UNDERSTANDING REGIONAL DIVERSITY IN THE HUMAN BILIARY TREE THROUGH TRANSCRIPTOMIC PROFILING OF PRIMARY TISSUES AND *IN VITRO* DERIVED ORGANOIDS

Casey Allison Rimland

Trinity College

Department of Surgery

University of Cambridge

This dissertation is submitted for the degree of Doctor of Philosophy

December 2018

This dissertation is dedicated to the 30 organ donors and their families who selflessly donated their organs making this research possible.

I will strive to honour your gift each day in my future career as both a physician and a scientist. Thank You.

DECLARATION

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as specified in the text.

This dissertation is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution.

In accordance with the guidelines of the degree committee of Clinical Medicine this dissertation does not exceed 60,000 words and it contains less than 150 figures.

Signed: 

Date: 26/11/2018

Casey Allison Rimland, Bachelor of Science

SUMMARY

Understanding regional diversity in the human biliary tree through transcriptomic profiling of primary tissues and *in vitro* derived organoids

Casey Allison Rimland

The biliary tree is a series of ductular tissues responsible for the drainage of bile produced by the liver and pancreatic secretions from the pancreas. The biliary tree is affected by a diversity of life-threatening diseases collectively called cholangiopathies. Cholangiopathies show regionalization, with some diseases such as biliary atresia predominantly targeting extrahepatic bile ducts (EHBDs) outside of the liver. Despite this, little is known on whether anatomical location within the biliary tree contributes to differences in functionality of biliary epithelium, especially in the EHBD compartment. Additionally, reports have demonstrated the possibility for *in vitro* culture of bile duct stem/progenitor cell organoids from both intrahepatic (IHBD) and EHBD sources. The relation of these organoid systems to each other, and to their tissue of origin, is largely unknown.

In this dissertation, I address these major questions by combining transcriptional analyses and *in vitro* culture of human bile duct organoids derived from primary IHBD and EHBD epithelium. First, I show that *in vitro* organoids can be derived from four regions of the human biliary tree: gallbladder, common bile duct, pancreatic duct, and intrahepatic bile ducts. Characterization of these organoids demonstrated expression of adult stem cell (LGR5/PROM1) and ductal (KRT19/KRT7) markers suggesting these cultures contained cells with a biliary stem/progenitor phenotype. Further, I show that IHBD organoids are distinct from EHBD organoids requiring different conditions for sustained growth. Using RNA-Sequencing, I demonstrate that primary tissues from different regions of the extrahepatic biliary tree display unique expression profiles and identify novel tissue-specific markers. I also show that only a limited number of these tissue specific differences are maintained in the *in vitro* organoids and that the organoids are very different from their tissue of origin. Finally, I demonstrate that IHBD, but not EHBD organoids, express a low-level of hepatocyte-specific markers under differentiation conditions.

Taken together, the work in this dissertation has uncovered regional specific markers for different anatomical regions of the human biliary tree. Further, I demonstrate that major differences exist between IHBD organoids and EHBD organoids *in vitro* and discover that only IHBD organoids have the capacity to express hepatocyte markers under differentiation conditions. Ultimately, these results may help to identify new targets for therapeutic development for cholangiopathies and regenerative medicine. They have also provided important insight to the understanding of both basic biliary physiology and also the field of biliary stem/progenitor cell organoids.

ACKNOWLEDGEMENTS

This work would not have been possible without a large number of people whom I have had the privilege to work with over the past four years. Firstly, I would like to thank my co-supervisors Ludovic Vallier and Tom Wynn. Thank you both for giving me the opportunity to work in your laboratories and pursue this project. The examples you have both set will be remembered throughout my career and I know my future as a physician scientist will be shaped strongly by the time I spent in both of your labs.

To all of the Cambridge LRMers, in particular Anna, Loukia, Elisa, Miguel, Carola, Rute, Samantha, Brandon, Imbisaat, Daniel, Will, Alexander, thank you for giving me an incredible support system while I was at Cambridge. From our lunch "field trips" to Collin the Caterpillar tea times, I never felt alone and knew every single one of you had my back. I could not have asked for a better group of colleagues to begin my PhD with.

To Nick, thank you for basically holding my hand for six months and teaching me all of the skills I needed to succeed as a stem cell biologist. Your supervision and guidance during the first year of my PhD was invaluable and even three years later I still hear your voice in my head when I do my tissue culture.

To Simone, the best summer student anyone could ever hope to have, thank you for reminding me of why I pursued a PhD in the first place. Being a mentor to you and seeing you grow into an independent scientist is one of the proudest accomplishments of my PhD. Plus, because of you, I will never be able to load a 384-well PCR plate without our animal game!

To the Wynn lab, a.k.a the Wynners, thank you for making my transition to the NIH a welcoming one. Especially to Kayla, thank you for being my friend and mentor when I so desperately needed one, thank you for the hugs, and thank you for sticking with me to the bitter end as the Wynn lab closed down.

To all of the many scientific collaborators who made this dissertation possible, especially Dr. Kourosh Saeb-Parsy, Dr. Wei-Yu Lu, Dr. Stuart Forbes, and Mr. Nikitas Georgakopoulos, thank you for giving so freely of your time and expertise to help me achieve this project.

To the Gates Cambridge Scholars program, thank you for not only funding the first two years of my PhD work, but for also providing me with the most vibrant and enriching community I have ever been a part of. I am so proud to call myself a Gates Scholar and

am still honoured to have been chosen to be a part of this community. Know that I will strive to go forth and make the world a better place!

To my Gates family members: Paulo and Juliana, Michael and Jess, Tercia and Phil, thank you for making Cambridge our home away from home and now for giving us family all around the world!

To Trinity College, thank you for providing me with both an academic and social home for the two years I was in Cambridge. Also, thank you for providing the fairy-tale backdrop to my Wedding.

To the Rheumatology department in NIAMS at the NIH, thank you for letting me spend time seeing patients and keeping up my clinical skills during my PhD work. In particular, to Dr. James Katz, thank you for taking me on as your medical student. I could not have asked for a better clinical mentor and teacher.

To my parents, thank you for everything you have done for me. I would not be here today without you both. I love you.

Lastly, to Aaron, my soulmate, best friend, and husband, thank you for standing by me throughout this roller coaster adventure of mine. None of this could have been possible without you and the countless sacrifices you made and continue to make in order for me to follow my dreams. I love you so much.

TABLE OF CONTENTS

LIST OF TABLES

List of Figures

LIST OF ABBREVIATIONS AND ACRONYMS

2D	Two-dimensional
3D	Three-dimensional
ABCA1	ATP Binding Cassette Subfamily A Member 1
ABCB1	ATP Binding Cassette Subfamily B Member 1
ABCC2	ATP Binding Cassette Subfamily C Member 2
ABCC3	ATP Binding Cassette Subfamily C Member 3
ABCG2	ATP Binding Cassette Subfamily G Member 2
ABCG5/8	ATP Binding Cassette Subfamily G Member 5/8
AC	Adenylate Cyclase
Ach	Acetylcholine
adj p-value	False Discovery Rate
adSCs	Adult Stem Cells
ADV/DMEM F12	Advanced Dulbecco's Modified Eagle Medium/Ham's F-12
AE2	Anion Exchanger of Type 2
AFP	Alpha Fetoprotein
ALB	Albumin
AMP	Adenosine Monophosphate
ANOVA	Analysis of Variance
APC	Adenomatous Polyposis Coli
APOA1	Apolipoprotein A-I
AQP	Aquaporin
ASBT/SLC10A2	Apical Sodium Bile Acid Transporter
ATP	Adenosine Triphosphate
BA	Bile Acids
BMP	Bone Morphogenic Protein
bp	Base Pairs
BSA	Bovine Serum Albumin
°C	Degrees Celsius
Ca^{2+}	Calcium Ion
CA4	Carbonic Anhydrase 4
CASP1	Caspase 1
CBCs	Crypt Based Columnar Cells

CBD	Common Bile Duct
CCK	Cholecystokinin
CCl_4	Carbon Tetrachloride
CD24	Cluster of Differentiation 24
CDK6	Cyclin-dependent Kinase 6
cDNA	Complementary DNA
CDX2	Caudal Type Homeobox 2
CFTR	Cystic Fibrosis Transmembrane Receptor
CHST4	Carbohydrate Sulfotransferase 4
Cl^-	Chloride Ion
cm	Centimeter
CYP3A4	Cytochrome P450 Family 3 Subfamily A Member 4
CYP7B1	Cytochrome P450 Family 7 Subfamily B Member 1
DAPI	4′,6-diamidino-2-phenylindole
DAPM	4,4’-methylene dianiline
DEXA	Dexamethasone
DKK-1	Dickkopf WNT Signalling Pathway Inhibitor 1
DM	Differentiation Media
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic Acid
ECM	Extracellular Matrix
ECOs	Extrahepatic Cholangiocyte Organoids
EGF	Epidermal Growth Factor
EHBD	Extrahepatic Bile Duct
EM	Expansion Media
EPCAM	Epithelial Cell Adhesion Molecule
ES Cells	Embryonic Stem Cells
ESLD	End-stage Liver Disease
ET-1	Endothelin-1
FAH	Fumarylacetoacetate Hydrolase
FCS	Fetal Calf Serum
FGF	Fibroblast Growth Factor
FMO	Fluorescence Minus One
FOXL1	Forkhead Box L1
FSK	Forskolin
FXR	Farnesoid X Receptor (also called NR1H4)

FZD	Frizzled
GBD	Gallbladder
GGT	Gamma-Glutamyl Transferase
GO	Gene Ontology
GSK3B	Glycogen Synthase Kinase-3 Beta
H^+	Hydrogen Ion
HBSS	Hank's Balanced Salt Solution
HCO_3^-	Bicarbonate Ion
HES1	Hairy and Enhancer of Split 1
hESCs	Human Embryonic Stem Cells
HEX	Haematopoietically Expressed Homeobox
HGF	Hepatocyte Growth Factor
HNF1A	Hepatocyte Nuclear Factor 1 Alpha
HNF1B	Hepatocyte Nuclear Factor 1 Beta
HNF4A	Hepatocyte Nuclear Factor 4 Alpha
HNF6	Hepatocyte Nuclear Factor 6 (also called Onecut 1)
HOXB2	Homeobox B2
hr	Hour
HSCs	Hematopoietic Stem Cell
HybHeps	Hybrid Hepatocytes
ICM	Inner Cell Mass
IF	Immunofluorescence
IgG_4	Immunoglobulin G4
IHBD	Intrahepatic Bile Duct
IHH	Indian Hedgehog
iPS Cells	Induced Pluripotent Stem Cells
K^+	Potassium Ion
KI67	Proliferation Marker Protein Ki-67
KRT19	Keratin 19
KRT7	Keratin 7
LGR5	Leucine Rich Repeat Containing G Protein-Coupled Receptor 5
LPCs	Liver Progenitor Cells
LRP5/6	LDL Receptor Related Protein 5/6
LSM700	Laser Scanning Microscope 700
mESCs	Mouse Embryonic Stem Cells

Mg^{2+}	Magnesium Ion
min	Minutes
mL	Milliliter
mM	Millimolar
MOGAT1	Monoacylglycerol O-Acyltransferase 1
mRNA	Messenger Ribonucleic Acid
MSCs	Mesenchymal Stem Cells
MUC1	Mucin 1
Na^{+}	Sodium Ion
NAC	N-acetylcysteine
NaCl	Sodium Chloride
NANOG	Homeobox Transcription Factor Nanog
ng	Nanogram
NHE3	Sodium Hydrogen Exchanger 3
NHSBT	National Health Service Blood and Transplant
NIAID	National Institute of Allergy and Infectious Diseases
NIC	Nicotinamide
NIH	National Institutes of Health
OATP-A	Solute Carrier Organic Anion Transporter Family Member 1A2 (also called SLCO1A2)
OCT	Optimal Cutting Temperature
OCT4	Octamer-binding Transcription Factor 4
OSM	Oncostatin-M
OSTA/B	Organic Solute Transporter Subunit Alpha/Beta
PancD	Main Pancreatic Duct
PBGs	Peribiliary Glands
PBS	Phosphate Buffered Saline
PCA	Principal Component Analysis
PCNA	Proliferating Cell Nuclear Antigen
PDGs	Pancreatic Duct Glands
PDX1	Pancreatic and Duodenal Homeobox 1
PFA	Paraformaldehyde
PGE2	Prostaglandin E2
PH	Primary Human Hepatocytes
PKA	Protein Kinase A
PPARG	Peroxisome Proliferator Activated Receptor Gamma

PPY	Pancreatic Polypeptide
PROM1	Prominin 1
PXR	Pregnane X Receptor (also called NR1I2)
qPCR	Quantitative Polymerase Chain Reaction
RIN	RNA Integrity Number
RNA	Ribonucleic Acid
RNA-Seq	RNA-Sequencing
RNF43	Ring Finger Protein 43
RRV	Rhesus Rotavirus
RSPO	R-Spondin
S1P2	Sphingosine 1-Phosphate Receptor Type 2
SC	Sigmoid Colon Organoids
SCTR	Secretin Receptor
SEM	Standard Error of the Mean
SHH	Sonic Hedgehog
SOX17	Sex Determining Region Y-Box 17
SOX2	Sex Determining Region Y-Box 2
SOX9	Sex Determining Region Y-Box 9
SPP1	Secreted Phosphoprotein 1
SST	Somatostatin
TA	Transit-Amplifying
TBX3	T-Box 3
TCF/LEF	T-cell Factor/Lymphoid Enhancer Factor
TGF-ß	Transforming Growth Factor Beta
TGR5	G-Protein Coupled Bile Acid Receptor 1
TTR	Transthyretin
U	Units
UBC	Ubiquitin
UGT	Uridine 5'-diphospho-glucuronosyltransferase
VDR	Vitamin D Receptor
VIM	Vimentin
VIP	Vasoactive Intestinal Polypeptide
VIPR1	Vasoactive Intestinal Polypeptide Receptor 1
WNT	Wingless-Type
x g	Times Gravity
µg	Microgram

µL	Microliter
µm	Micrometer
µM	Micromolar

1 INTRODUCTION

1.1 Characteristics of Mammalian Stem Cells

Stem cells are unique cells found in multicellular organisms which possess two key characteristics (**Figure 1.1**).[1,2] The first is the capacity for self-renewal, the ability to divide and give rise to additional stem cells. The second is the capacity for differentiation, in which stem cells must also be able to divide under certain conditions to yield differentiated cells with specialized functions. Stem cells are broadly classified by their source and/or potency (**Table 1.1**). Potency, or plasticity, refers to the range of cells a stem cell is capable of producing.[2] Totipotent stem cells are capable of giving rise to all three germ layer cell types (endoderm, mesoderm, and ectoderm) of an organism as well as extra-embryonic tissues. Pluripotent stem cells are more restricted and cannot give rise to extra-embryonic tissues. Multipotent stem cells are even more restricted and are only able to give rise to a select few types of specialized cells found within a tissue. Lastly, unipotent stem cells only give rise to a single specialized type of cell while also having the capacity for self-renewal.

Classification		Definition	Examples
Source/Type			
	Embryonic Stem Cells (ES Cells)	Pluripotent stem cells derived from the inner cell mass of the blastocyst embryo	Human ES cell lines
	Induced Pluripotent Stem Cells (iPS Cells)	Pluripotent stem cells derived by direct epigenetic reprogramming of somatic cells through ectopic expression of specific pluripotency factor genes	Human iPS cell lines
	Adult Stem Cells (adSCs)	Stem cells which are found in adult tissues. These cells tend to be restricted in their differentiation potential and only differentiate into tissue-specific specialized cells. These cells are important in tissue repair and regeneration.	Intestinal Lgr5+ stem cells, hematopoietic stem cells (HSCs), mesenchymal stem cells (MSCs)
Potency			
	Totipotent	Cells which have the potential to give rise to all cell types in an organism including embryonic tissues from all three germ layers (endoderm, mesoderm, ectoderm) as well as extra-embryonic trophectoderm placental tissue.	Zygote to Morula
	Pluripotent	Cells which have the potential to give rise to all cell types of the embryo but cannot give rise to extra-embryonic placental tissues.	Human ES Cells from the blastocyst inner cell mass
	Multipotent	Cells which have a limited differentiation potential confined to producing only differentiated cells found in a single tissue type	Intestinal LGR5+ stem cells, HSCs/MSCs
	Unipotent	Cells which retain the ability of self-renewal and proliferation, however they are capable of producing only one differentiated cell type.	Muscle Satellite cells

Table 1.1 Classification of stem cells.

Table detailing the classification systems for different types of stem cells. Modified with permission from Hui et al. 2011.[2]

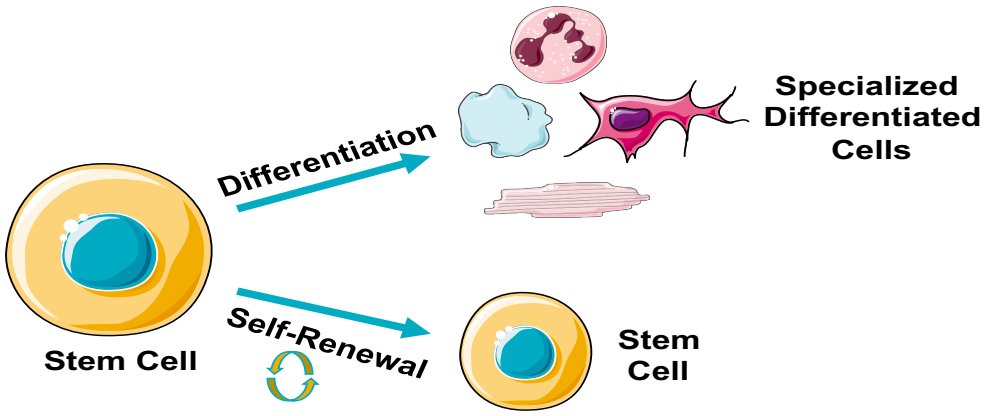


Figure 1.1 Stem Cell Biology.

Diagram depicting the two key characteristics, self-renewal and differentiation capacity, which are unique to stem cells. Figure created using artwork from Servier Medical Art under a creative commons license.

1.1.1 Embryonic and induced pluripotent stem cells

Embryonic stem cells are pluripotent cells derived from the inner cell mass of blastocysts.[2–4] Work on embryonic stem cells (ESCs) initially began in the mouse, with mouse ESCs (mESCs) being discovered in 1981.[5] Mouse ESCs were found to provide a source of pluripotent cells with essentially limitless proliferation capacity and were capable of *in vitro* differentiation to all three germ cell layers. The historical work on mESCs paved the way for the isolation and establishment of human ESCs (hESCs) in 1998 by James Thomson.[1,6] hESCs are derived from donated embryos which were originally produced for clinical *in vitro* fertilization therapy.[3,6] These embryos, at the 5-day blastocyst stage, are used to isolate inner cell mass (ICM) cells which are then used to establish *in vitro* hESC lines (**Figure 1.2**).[3] Several signalling pathways have been shown to be required for maintaining pluripotency of hESCs including the Activin/Nodal/TGF-ß and FGF pathways.[7] hESCs express several characteristic genes which are integral in maintaining pluripotency such as OCT4, NANOG, and SOX2.[8,9]

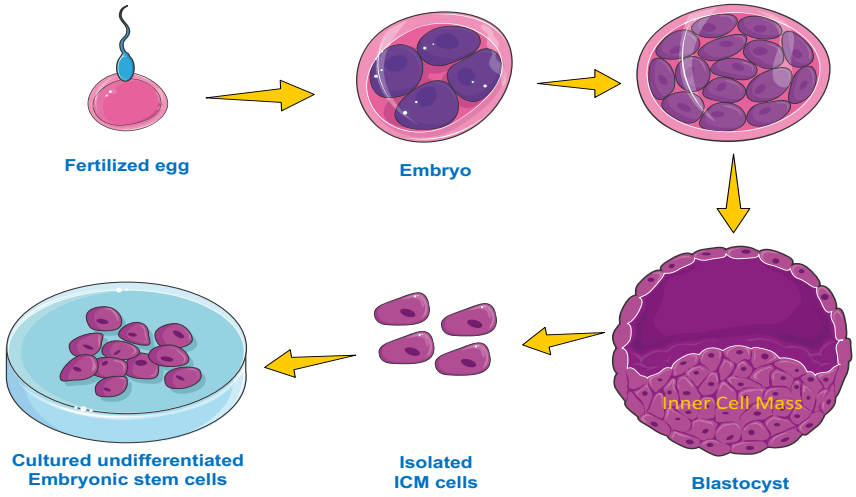


Figure 1.2 Origin and derivation of ESCs.

Diagram depicting the developmental origin and derivation of ESCs. Figure created using artwork from Servier Medical Art under a creative commons license.

Induced pluripotent stem cells (iPSCs) were first described by Shinya Yamanaka's group in 2006 in mice and in 2007 with human cells.[10,11] iPSCs are pluripotent cells initially derived from mature, somatic cells which express a similar transcriptional profile and show similar differentiation capacity as ESCs. iPSCs are produced by taking a somatic cell, for instance a fibroblast, and exposing the cell to four transcription factors such as OCT4, SOX2, KLF4, MYC. These factors were initially introduced into cells through integrational retrovirus/lentivirus gene delivery techniques.[12] However, recent advents in the field allow for non-integrative methods such as episomal DNA, direct mRNA delivery, or recombinant protein delivery, which eliminates the risk of mutations occurring at the site of integration.[12] The development of iPSC technology has revolutionized the field of stem cell biology. iPSCs allow for stem cell derivation from patients with rare diseases to develop, in most cases for the first time, an *in vitro* model of human disease. Additionally, the advent of iPSC technology provides hope for one day being able to derive allogenic stem cell lines from patients that can be then be used for cell based therapies and personalized medicine.[12]

Due to their extensive *in vitro* expansion and differentiation capacities, both ESCs and iPSCs have been proposed as cell sources for regenerative medicine and cell therapy. However, several challenges still exist with these cells that must be overcome before they become useful for clinical applications. This includes the lack of *in vitro* differentiation

protocols to produce mature, adult cell types. For example, current protocols for differentiating hepatocyte-like-cells from hESCs/iPSCs produce cells more closely resembling fetal hepatocytes than adult hepatocytes.[13–15] Additionally, the ethical implications of using hESCs are still of concern due to their embryo-derived nature. However, the availability of iPSCs may help to bypass this ethical problem and may serve as an alternative pluripotent cell source. Finally, the potential for teratoma formation by contaminating pluripotent cells in a sample of iPSC/ESC derived differentiated cells is another concern when considering these cells for therapy.

1.1.2 Adult tissue stem/progenitor cells

Adult stem cells (adSCs) are multipotent stem cells which reside in adult tissues and are capable of self-renewal and differentiation into a limited number of organ/tissue-specific cell types.[2,16,17] Historically, the *gold-standard* adSC was hematopoietic stem cells (HSCs).[18] HSCs are rare cells, found in the bone marrow, which are capable of self-renewal. They are also able to differentiate into all blood cell lineages. HSCs divide to give rise to progressively more differentiated progenitor cell populations. These progenitor cells then further differentiate to give rise to mature blood cell lineages including red cells, megakaryocytes, myeloid cells, and lymphocytes. The presence of HSCs is what allows for bone marrow transplantation to be an effective therapy and is the first example of stem cell therapy in humans.[18]

Recently, an additional source of adSCs has been described. These adSCs have been discovered in a remarkably large number of epithelial lined tissues including the skin, stomach, intestine, mammary glands, kidney, prostate, and others.[17,19–24] These epithelial adSCs, despite residing in specialized tissues with unique functional requirements, have been found to rely on similar niche signals across tissues.[17,25] One pathway that is important for maintaining these stem/progenitor cells is the WNT signalling pathway. This pathway is well-conserved in multicellular organisms.[25–27] WNT proteins are important regulators of development and morphogenesis in multiple tissues and there are over 20 WNT proteins in mammals. The WNT signalling pathway is complex with several different divergent pathways within it. The most well-known is the canonical WNT pathway (**Figure 1.3**).

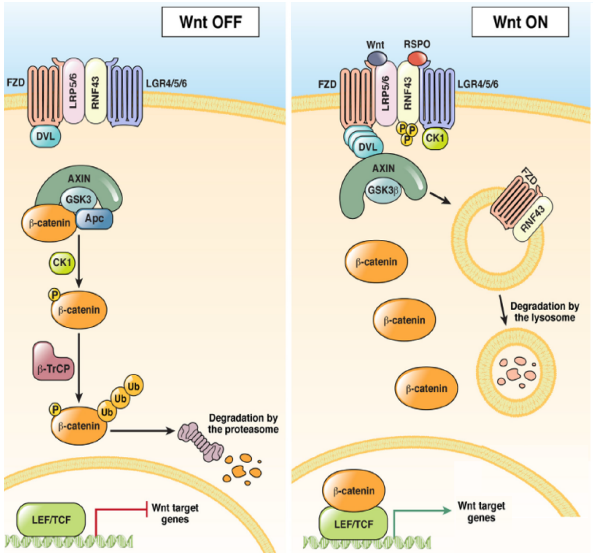


Figure 1.3 Canonical WNT signalling pathway and the LGR5/RSPO axis.

Diagram depicting the canonical WNT signalling pathway and its amplification through LGR5/RSPO. Modified with permission from Koo et al (2014).[28]

This pathway consists of WNT ligands that bind to a family of Frizzled (FZD) receptors. This initiates a signalling cascade resulting in the activation of ß-catenin which translates into the nucleus to activate the TCF/LEF transcription factor. This ultimately leads to the transcription and translation of WNT target genes which exert the biological effects of the WNT pathway.[28] When WNT ligands are not present, ß-Catenin is inactivated through phosphorylation by the destruction complex (members include APC, Axin, GSK3ß). Phosphorylated ß-catenin is targeted to the proteasome for degradation. The canonical WNT pathway has been shown to be critical in maintaining self-renewal and proliferation of epithelial stem cells. Further, the orphan G-protein coupled receptor LGR5 has recently been identified as a marker of adult epithelial stem cells in multiple tissues. LGR5, along with its ligands R-Spondin 1, 2, and 3 (RSPO), acts to modulate and enhance WNT signalling by inhibiting RNF43.[28] RNF43 usually targets the FZD and LRP5/6 receptors for degradation and acts as a negative regulator of WNT signalling. LGR5/RSPO prevent this, thereby increasing the longevity of FZD and LRP5/6 receptors on the cell surface potentiating WNT signalling (**Figure 1.3**). The discovery of the LGR5/RSPO axis has been critical for the isolation of LGR5+ adSCs from a number of epithelial tissues and the discovery of RSPO has made it possible to culture these cells *in vitro* using recombinant RSPO proteins.

1.2 Hepatobiliary compartment – Liver

The liver is the largest internal organ in the human body, weighing over one kilogram and making up 2.5% of adult body weight.[29,30] The liver is responsible for hundreds of critical biological functions such as the detoxification of xenobiotics and metabolic breakdown products, lipid and cholesterol synthesis, red blood cell recycling, protein synthesis including blood clotting factors and complement, and glycogen storage.[31] The liver is macroscopically divided into two lobes and receives blood supply from both the hepatic artery and portal vein, with vascular drainage accomplished via the central vein.[29–31]

Human diseases of the liver are responsible for high morbidity and mortality and are the 12th leading cause of death in the United States.[32] Cirrhosis is a common endpoint for patients with end-stage liver disease (ESLD). Common complications of cirrhosis include portal hypertension, ascites, hepatic encephalopathy, variceal haemorrhaging, and death.[32] Cirrhosis develops after years of chronic injury to the liver that results in progressive fibrosis and loss of functional liver tissue.[33] Aetiologies of liver fibrosis are varied and include those such as alcoholic liver disease, *Schistosoma* infection, viral hepatitis, non-alcoholic fatty liver disease, autoimmune hepatitis, a variety of genetic disorders including inborn errors of metabolism, and drug or toxin induced liver failure.[33] Currently, ESLD can only be cured by orthotopic liver transplantation.[32,34]

1.2.1 Anatomy and embryology of the liver

At the microscopic level, the adult liver is highly organized. There are five specific cell types of the liver well characterized so far (**Table 1.2**).[31] These cell types are classified into either parenchymal or non-parenchymal types. Parenchymal cells consist of hepatocytes which are the main functional cell type of the liver. Non-parenchymal cells consist of hepatic stellate cells, kupffer cells (resident macrophages), sinusoidal endothelial cells, and cholangiocytes. The basic organizational unit of the liver is the hepatic lobule, composed of a central vein surrounded by portal triads **(Figure 1.4)**.

Cell Type	Function
Hepatocytes	~70% of the liver cells Protein synthesis and secretion Bile acid synthesis and secretion Cholesterol metabolism Detoxification Urea metabolism Glucose/Glycogen metabolism Acute phase response Blood clotting
Cholangiocytes	~3% of the liver cells Form ductular system to transport bile Control bile flow Modify and concentrate bile Secrete H_20 and Bicarbonate Control pH of bile
Stellate Cells	~1-2% of liver cells Extracellular matrix maintanence Activated to become myofibroblasts Participate in regenerative response to injury Vitamin A storage Secrete cytokines
Endothelial Cells	~2.5% of liver cells Vasculature cells Form veins, arteries, venuoles, and arterioles Control blood flow Paticipate in parenchymal zonation
Kupfer Cells	~2% of liver cells Resident macrophages Scavengers of foreign material

Table 1.2 Predominant cell types of the liver and their functions.

Modified with permission from Si-Tayeb et al (2010).[31]

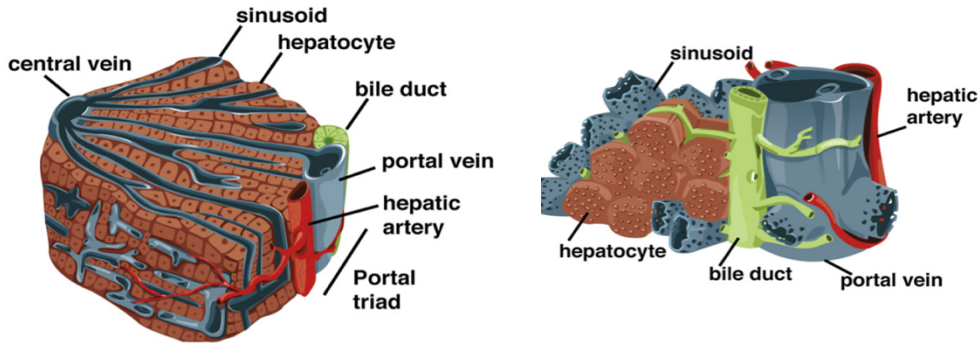


Figure 1.4 Functional anatomy and microscopic architecture of the liver.

Diagram depicting the microscopic anatomy of the liver lobule and the components of the portal triad. Image modified with permission from Si-Tayeb et al (2010).[31]

Portal triads consist of a portal vein, hepatic artery, and bile duct. Blood flows from the portal vein to the central vein via sinusoids. Hepatocytes lining sinusoids uptake, modify, and transport substances actively into the circulation. Hepatocytes also produce bile, which is secreted into bile canaliculi that are lined with cholangiocyte cells. Bile ultimately drains into larger ducts and out of the liver through the biliary system. This fully mature liver structure is not obtained until the post-natal period.[31] During embryogenesis, the liver is derived from the definitive endoderm and ventral foregut.[31,35] Hepatogenesis relies heavily on reciprocal signalling interactions between the endoderm and mesoderm. The hepatic diverticulum forms from the ventral foregut and is the first evidence of liver development. The hepatic diverticulum is in close proximity to the embryonic heart and gives rise to the liver bud. The liver bud is composed of bipotential progenitors, called hepatoblasts, which differentiate to form hepatocytes and intrahepatic bile duct cholangiocytes. The septum transversum mesenchyme gives rise to stellate cells in the liver. As the embryonic liver continues to develop, it is vascularized and ultimately acts as the primary site of embryonic haematopoiesis.[31,35]

1.2.2 Liver regeneration and homeostatic turnover

The liver has long been known to possess a remarkable, yet complex capacity for regeneration.[36] After partial hepatectomy, rodents are capable of regenerating their entire pre-surgical liver mass in as little as 5-7 days.[36,37] Liver regeneration and the predominant mechanism by which new cells are derived, is dependent on the type of liver injury pattern present. In the partial hepatectomy model, it has been shown that regeneration is driven

largely by the proliferation of mature cell types in the liver, which re-enter the cell cycle and divide to supply new hepatocytes and restore tissue mass.[36,37]

However, in injury models where mature hepatocyte growth is impaired, such as in chronic fibrotic disease, it has been suggested that liver stem/progenitor cells (LPCs) may become the primary source of cells for regeneration, although the origin and identity of these cells is still controversial in the field.[36–40] It has been suggested that both the hepatocyte and biliary compartments of the liver can act as sources of regenerating cells. Which compartment becomes activated may depend on the degree and/or aetiology of liver injury. In the hepatocyte compartment, two sub-populations of hepatocytes are thought to contribute to liver regeneration under different circumstances: (1) Pericentral diploid, Axin-2 positive hepatocytes contribute to homeostatic liver turnover and (2) Periportal Hybrid Hepatocytes (HybHeps) participate in regeneration under states of chronic liver injury.[39,41] Within the biliary compartment, it has been proposed that a subpopulation of intrahepatic biliary epithelial cells, called oval cells or liver progenitor cells, have the ability to differentiate into both cholangiocytes and hepatocytes.[42–44] Oval cells are thought to act as the regenerative cell type when hepatocyte division is impaired.[38,45] These biliary progenitor cells will be discussed in more detail below in relation to the stem cells of the biliary compartment. Here we will focus on the hepatocyte compartment.

1.2.2.1 Pericentral diploid axin2-positive hepatocytes

Wang et al (2015) first described a population of hepatocytes which lie adjacent to the central veins in murine livers.[41] In the course of their work, they noticed that a portion of hepatocytes lining central veins expressed the WNT pathway target gene AXIN2 along with the early hepatoblast/pluripotency gene TBX3. Using lineage-tracing techniques, they were able to demonstrate that Axin2 positive hepatocytes were capable of self-renewal while also giving rise to 40% of all hepatocytes in the liver over the course of one year. These pericentral hepatocytes were also found to be unique in that they were diploid, whereas most mature hepatocytes are polyploid. Polyploid cells are thought to have a lower replicative potential and are usually associated with senescent cells in comparison to diploid cells. The authors suggest that the diploid status of the Axin2 positive pericentral hepatocytes may explain their ability to continue to divide and restore hepatocytes under homeostatic conditions. The authors also demonstrated that the central

vein endothelium secretes WNT ligands (WNT2 and WNT9b) thought to be the source of WNT signals responsible for maintaining the niche of these pericentral hepatocytes. This work represents a major discovery into the regenerative mechanisms at play in the liver. Further, for the first time, they demonstrated that a population of cells, with stem cell like characteristics, exists in the hepatocyte compartment of the liver and are responsible for hepatocyte turnover in homeostasis. The role these pericentral hepatocytes play in liver injury was not assessed and remains to be studied. In 2015, it was discovered that an additional population of hepatocytes, located at the opposite end of the liver lobule at the portal vein, are capable of regenerating the damaged liver. These cells, called periportal hybrid hepatocytes (HybHPs), are discussed below.

1.2.2.2 Periportal hybrid hepatocytes

HybHPs were first described by Font-Burgada, et al (2015).[39] Their work identified a population of hepatocytes adjacent to the portal vein and in close contact to the bile ducts. These periportal hepatocytes were named hybrid hepatocytes due to their expression of both biliary and hepatocyte lineage markers (**Figure 1.5**). Using lineage-tracing techniques these HybHPs were shown to proliferate in mice following treatment with liver toxins such as carbon tetrachloride. Further, HybHPs showed high engraftment ability in a $FAH^{-/-}$ mouse model. These HybHPs were transcriptionally unique from conventional hepatocytes and bile ducts. However, the HybHPs expressed several key markers of biliary cells including SOX9 and SPP1, while also expressing markers of mature hepatocytes such as HNF4A and HNF1A. Interestingly, expression of oxidative metabolism genes was lower in HybHPs than conventional hepatocytes. The authors suggest that this may mean HybHPs do not play a large role in the detoxification function of the liver. HybHPs may then be protected from many of the insults related to liver toxicities, making them a likely cell type to participate in regeneration.[39]

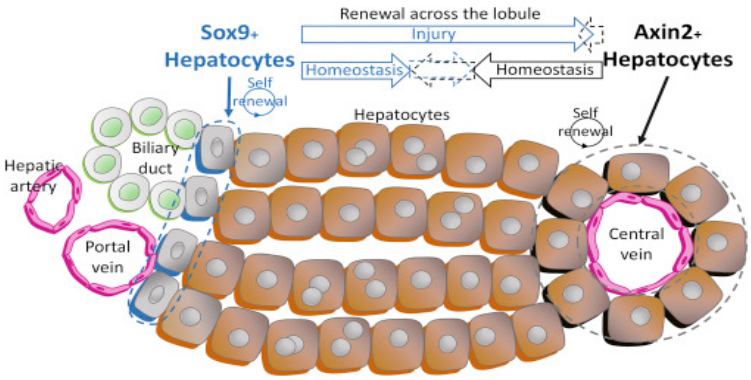


Figure 1.5 Diagram depicting the two sources of regenerative hepatocytes in the liver.

Reproduced with permission from Bird et al (2015).[46]

1.3 Hepatobiliary Compartment – Biliary Tree

1.3.1 Anatomy of the biliary and pancreatic duct systems

The biliary tree is a series of ductular tissues whose primary responsibilities include the drainage and modification of bile produced by the liver and pancreatic digestive secretions from the pancreas.[47,48] The biliary system can be anatomically divided into three broad compartments: intrahepatic, extrahepatic, and pancreatic (**Figure 1.6**). The intrahepatic compartment, which consists of the bile ducts inside the liver parenchyma, is further divided based on duct diameter.[49] The smallest intrahepatic ducts begin at the canals of Hering and proceed in increasing size from terminal cholangioles (<15 um diameter), interlobular ducts (15-100 μm), septal ducts (100-300 μm), area ducts (300-400 μm), segmental ducts (400-800 μm) to ultimately terminate at the hepatic ducts (>800 μm).[49] The right and left hepatic ducts merge to form the common hepatic duct, which marks the beginning of the extrahepatic compartment. The extrahepatic compartment is defined not by duct size but instead by anatomical region. The extrahepatic compartment includes the common hepatic duct, gallbladder (GBD), cystic duct, and common bile duct (CBD). Additionally, inside the pancreatic parenchyma, a pancreatic ductal system also exists, which, similarly to intrahepatic bile ducts inside the liver, are categorized based on duct size.[50] Pancreatic duct diameter increases successively from the smallest intercalated ducts, to intralobular ducts, interlobular ducts, and ultimately merge to form the main pancreatic duct (PancD).[50] Common to each of these regions of the biliary tree

is a lining by specialized epithelial cells called cholangiocytes.[49] Cholangiocytes are highly polarized cells with an apical and basolateral domain. They contain apical microvilli and have primary cilia. Cholangiocytes ubiquitously express the cytokeratins KRT19 and KRT7 as well as epithelial cell markers such as EPCAM.[49] In the smallest intrahepatic and intrapancreatic ducts, cholangiocytes are cuboidal in shape. As duct diameter increases, cholangiocytes become columnar in shape. The large intrahepatic and intrapancreatic ducts, extrahepatic ducts, main pancreatic duct, and gallbladder are all lined by a simple columnar epithelium (**Figure 1.6**). Unique to the gallbladder is a folded mucosa with a structure resembling the intestine with an underlying smooth muscle layer allowing for gallbladder contraction.[51]

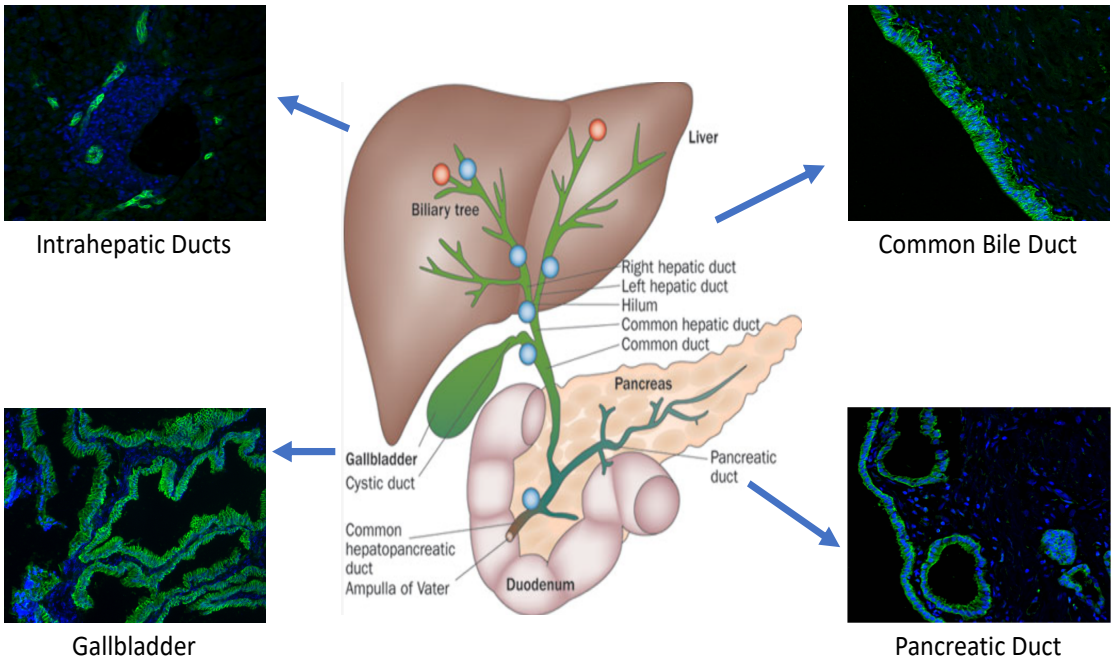


Figure 1.6 Anatomy and histology of the human biliary tree.

Diagram depicting the anatomy of the biliary tree (modified with permission from Cardinale et al. (2012)) and histological images of different regions of the biliary tree stained for EPCAM (green) an epithelial cell specific marker, with DAPI (blue) counterstaining of the nuclei.[52]

Additionally, the larger intrahepatic bile ducts, extrahepatic bile ducts and pancreatic ducts have glandular structures underlying the main epithelium. These glandular structures are called peribiliary glands (PBGs) or pancreatic duct glands (PDGs) in bile and pancreatic ducts respectively, and largely serve an unknown biological function.

Recently, the PBGs/PDGs have been shown to be connected to the main epithelium through small channels and are thought to harbour mucin producing cells.[53] Reports suggest that the PBGs/PDGs possess cells with a regenerative capacity and contain biliary stem/progenitor cells, however this still remains to be definitely shown.[54]

1.3.2 Heterogeneity in the intrahepatic biliary epithelium

Within the intrahepatic biliary tree, cholangiocyte heterogeneity exists.[55–58] Two types of cholangiocytes have been described: small and large. Small cholangiocytes, which are around 8-9 μm in diameter, line the smallest intrahepatic ductules and canals of Hering.[55] Large cholangiocytes, around 14-15 μm in diameter, line larger intrahepatic ducts.[55] Almost all of this work has been performed in rodents, led by Gianfranco Alpini and others. Using laser microdissection techniques, Alpini and colleagues were able to isolate intrahepatic bile duct units from rat livers and select for duct size, in order to specifically study small or large ducts independently.[59] Additionally, *in vitro* small and large cholangiocyte cell lines have been established using counterflow elutriation and immunoaffinity separation.[60] Subsequent investigation of these cell lines and also on the intrahepatic bile duct units, uncovered differential expression of markers and differences in biologic function between the two types of cholangiocytes (**Figure 1.7**).[55–60] Large cholangiocytes are responsible for biliary secretory functions and respond to secretin and somatostatin. The biological function of small cholangiocytes is less well-understood but they may act as a reservoir for regenerating large cholangiocytes in states of injury.[55–60] Ultimately this work indicates that cholangiocytes are not a uniform cell population within the liver. It is important to note that most of this work has been performed on rodent intrahepatic bile ducts and only limited work has been performed in human intrahepatic bile ducts.[56] Furthermore, although it is largely assumed that the large extrahepatic bile ducts are lined by cholangiocytes similar to the large intrahepatic cholangiocytes, the direct comparison of extrahepatic cholangiocytes to either type of intrahepatic cholangiocyte has not been performed.[61]

Markers	Small ducts	Large ducts	Function
γ-GT	Not expressed	Interlobular ducts	Glutathione metabolism
ALP	Not expressed	Interlobular ducts	Inhibition of secretin-induced choleresis
Cytochrome P4502E1	Not expressed	Expressed by large ducts	Dehalogenation of CCl_4
Lipase, α-amylase and trypsin	Human septal ducts	Large ducts, and peribiliary glands	Biliary development
Bcl-2	Human small ductules	Not expressed	Anti-apoptotic
Secretin receptor	Not expressed. *De novo* expressed after large duct damage	Expressed by large rodent ducts	Modulation of bicarbonate secretion
CFTR	Not expressed	Expressed by large ducts	Modulation of Cl− secretion
Cl^-/HCO_3^- AE2	Not expressed	Human and rodent ducts	Regulation of HCO_3^- secretion
Somatostatin receptor ($SSTR_2$)	Not expressed	Expressed by large rat ducts	Inhibition of growth and secretin-choleresis
Gastrin/CCK-B receptors	Unknown	Expressed by rat ducts	Inhibition of growth and secretin-choleresis
Melatonin receptor	Not expressed	Expressed by large rat ducts	Inhibition of cholestasis by
D2 dopamine receptors	Not expressed	Expressed by large rat ducts	Inhibition of secretin-induced choleresis
M3 acetylcholine receptors	Undefined	Expressed by large rat ducts	Stimulation of secretin-choleresis
α1 adrenergic receptors	Expressed by small rat ducts	Expressed by large rat ducts	Stimulation of secretin-induced choleresis
Endothelin receptors (ETA and ETB)	Expressed by small rat ducts	Expressed by large rat ducts	Inhibition of secretin-induced choleresis
Serotonin receptors	Undefined	Expressed by large rat ducts	Regulation of biliary growth
NFAT2 and NFAT4	Expressed by small ducts	Low expression in large ducts	Stimulation of small biliary growth

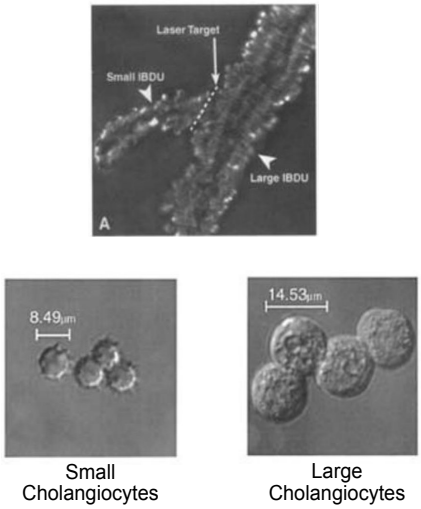


Figure 1.7 Characteristics of small and large intrahepatic cholangiocytes

Table detailing the differential expression of markers between small and large intrahepatic bile ducts as well as images showing a large and small intrahepatic bile duct unit and isolated small/large cholangiocyte cells. Figure modified with permission from Han et al (2013).[58] ALP: alkaline phosphatase; CFTR: cystic fibrosis transmembrane conductance regulator; $Cl^-/HCO3^-$ AE2: chloride bicarbonate anion exchanger 2; γ-GT: γ-glutamyl transpeptidase.

1.3.3 Function of the biliary and pancreatic duct systems

1.3.3.1 Transport and modification of bile

The main functional responsibility of the intrahepatic bile ducts, extrahepatic bile ducts, and gallbladder is to transport bile from the liver into the intestine. Bile is an aqueous solution made up predominantly of water (95%).[62,63] Bile production is initiated in the liver where hepatocytes actively secrete components across the canalicular membrane, which drain into the beginnings of the biliary tree, the canals of Hering. The liver can produce up to 500-1000 mL of bile a day. Bile is made up of the following major components:

(1) Bile Acids/Salts, (2) Phospholipids, (3) Cholesterol, (4) Bile pigments/Conjugated Bilirubin, (5) Inorganic Salts (K^+, Na^+, HCO_3^-), (6) Proteins, (7) Fatty Acids, (8) Xenobiotics such as drug metabolism by-products.[62,63]

Bile acids are 24-carbon molecules formed from the breakdown of cholesterol. Cholic Acid and Chenodeoxycholic acid, primary bile acids, are the two most commonly found in human bile.[62,63] Prior to secretion into bile, bile acids are conjugated to the amino acids, glycine or taurine, which increases their solubility and ability to form sodium salts. Once secreted, bile acids form micelles with phospholipids, of which phosphatidylcholine is the most common. These micelles help to solubilize free cholesterol in bile and allow bile secretion to be the predominant source of cholesterol excretion from the body. Additionally, bile acids are amphipathic molecules that promote fatty acid digestion once in the intestine. Bacteria in the intestine can also further modify primary bile acids to create secondary bile acids. Bile acids are almost entirely reabsorbed in the intestinal ileum by what is known as the enterohepatic circulation which returns bile acids directly into the portal venous circulation.[62,63] It has recently been shown that bile acids, beside playing a key role in cholesterol and lipid metabolism, are also capable of acting as signalling molecules through nuclear receptors FXR, PXR, and VDR as well as two G-protein coupled receptors TGR5 and S1P2.[64]

Bile pigments consist of conjugated bilirubin, the result of heme degradation pathways.[65] Bilirubin is formed through the oxidation of Heme, a by-product of red blood cell turnover. Heme is oxidized by the enzyme heme oxygenase to form biliverdin which is then converted to unconjugated bilirubin by the enzyme biliverdin reductase. Unconjugated bilirubin is transported to the liver in the blood stream while bound to albumin. Bilirubin is finally conjugated to glucuronic acid by members of the family of enzymes called uridine-diphosphoglucuronic glucuronosyltransferases. Bacterial flora in the colon can further modify conjugated bilirubin to urobilinogen and stercobilinogen which is excreted in the stool.[65]

Cholangiocytes of the intrahepatic and extrahepatic compartments actively participate in the modification of bile through both absorption and secretion, especially in the gallbladder.[51] These cholangiocytes have been shown to be capable of modifying the content of bile by the secretion or absorption of fluid, ions, lipids, cholesterol, bile acids, mucins, and even xenobiotic compounds.[51] These capabilities are possible due to the expression of various transporters on cholangiocytes that are summarized in **Figures 1.8 and 1.9**. Additionally, certain hormones and signalling molecules act on cholangiocytes

to promote fluid and bicarbonate secretion including secretin, vasoactive intestinal polypeptide (VIP), acetylcholine (ACh), and bombesin.[49] Other molecules act on cholangiocytes to inhibit secretion such as somatostatin (SST) and endothelin-1 (ET-1).[49] Many of these signalling factors are produced by intestinal neuroendocrine cells, coupling intestinal digestion states directly to bile modification. Interestingly, the expression of some transporters and receptors for signalling factors may differ along the biliary tree.[51]

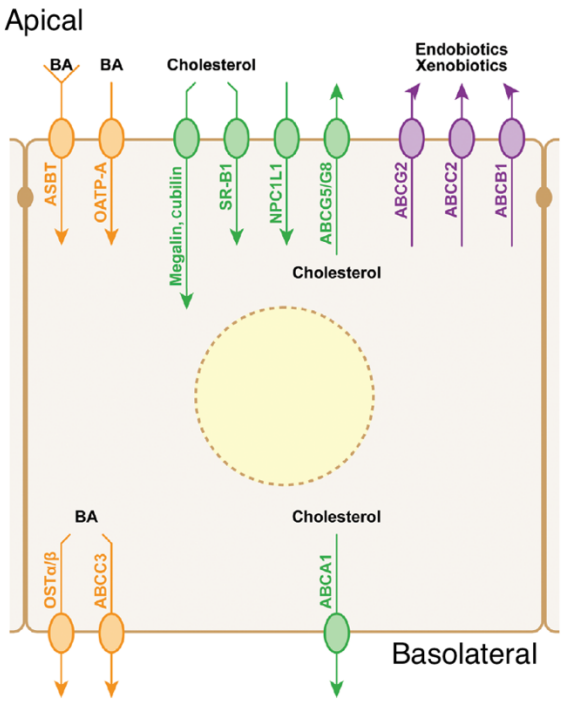


Figure 1.8 Diagram depicting the molecular mechanisms thought to regulate lipophilic molecule transport in bile duct cholangiocytes

Biliary cholangiocytes secrete and absorb various lipophilic molecules such as cholesterol, bile acids, and xenobiotics. BAs can be transported across the apical membrane of cholangiocytes through the Apical Sodium Bile Acid Transporter (ASBT/SLC10A2) or the OATP-A transporter. While bile acids efflux from the basolateral membrane through either the OST or ABCC3 transporters. Cholesterol transporters on the apical membrane include ones such as ABCG5/8. While cholesterol efflux is accomplished through the ABCA1 transporter on the basolateral membrane. Xenobiotics are secreted across the apical membrane through ABCG2, ABCC2, and ABCB1 transporters. Diagram modified with permission from Housset et al (2016).[51]

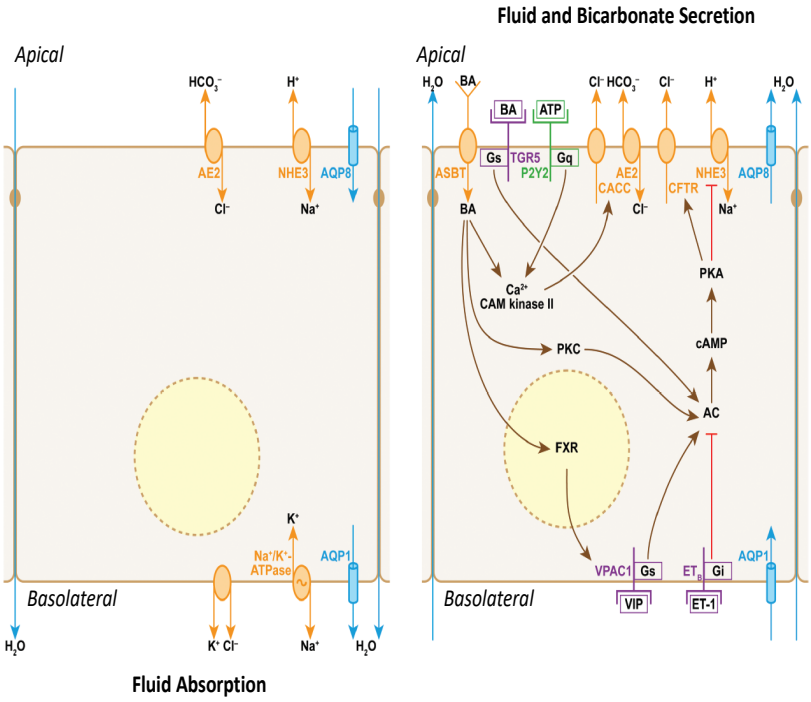


Figure 1.9 Diagram depicting the molecular mechanisms thought to regulate fluid and ion transport in cholangiocytes

On the left, isosmotic fluid absorption from the apical to basolateral side of cholangiocytes is depicted. NaCl is transported across the apical membrane through the Na^+/H^+ exchanger (NHE3) and the HCO_3^-/Cl^- exchanger (anion exchanger of type 2 or AE2). This allows sodium to enter cholangiocytes down its concentration gradient which is maintained by basolateral Na^+/K^+ ATPase active transport pumps. Water absorption then occurs passively down an osmotic gradient through paracellular transport or alternatively through aquaporin water transport channels (i.e. AQP1, AQP4, AQP8). On the right, molecular mechanisms involved in net fluid and bicarbonate secretion are depicted. Cyclic AMP (cAMP) is produced by Adenylate Cyclase (AC) which itself is modified by several pathways. Ultimately, cAMP acts to activate Protein Kinase A (PKA). PKA then inhibits the NHE3 transporter stopping fluid absorption. PKA also activates the Cystic Fibrosis Transmembrane Receptor (CFTR) which secretes chloride out of cholangiocytes, across the apical membrane, to create a chloride gradient which then shifts the AE2 exchanger in favour of net bicarbonate secretion. Additionally, bile acids (BA) and Vasoactive Intestinal Polypeptide (VIP) act to enhance these pathways. Water is then secreted passively down its osmotic gradient through AQPs or paracellular transport. Diagram modified with permission from Housset et al (2016).[51]

For example, secretin receptor (SCTR) was found to be expressed in intrahepatic bile ducts but not gallbladder epithelial cells, while the VIP receptor was highly expressed in gallbladder epithelial cells.[51,66] Gallbladder epithelial cells *in vitro* are responsive to secretin hormone treatment, but these experiments have usually been performed with supra-physiologic levels of secretin. Secretin has also been found to bind to VIP receptors

which may explain this *in vitro* observation.[51] Similarly, the apical sodium bile acid transporter (ASBT/SLC10A2) is thought to be highly expressed in gallbladder cholangiocytes but not intrahepatic cholangiocytes.[51] These observations have not been extensively studied and the relationship between different regions of the biliary tree are still not fully understood.

1.3.3.1 Transport and modification of pancreatic exocrine secretions

The pancreas, similarly to the liver, is a highly organized tissue that has two main functional compartments: endocrine and exocrine.[67] The exocrine compartment of the pancreas consists of exocrine acinar cells and ductal cholangiocytes and makes up over 95% of the total pancreatic mass.[68] The endocrine compartment, as well as stromal cells, nerves, and blood vessels, make up the remainder (**Figure 1.10**).[67,68] The endocrine, or hormone producing cells of the pancreas, reside in the Islets of Langerhans and are interspersed throughout the exocrine pancreas. Within the islets, specialized cells produce specific hormone types, many of which act to regulate blood glucose levels. Alpha-cells produce glucagon, Beta-cells produce insulin, Delta-cells produce Somatostatin, and PP-cells produce pancreatic polypeptide.[67,68]

The pancreatic ductal system within the pancreas makes up only 10% of the total tissue mass.[50] Pancreatic ducts, unlike bile ducts, which transport bile, function to transport pancreatic secretions to the intestine. Pancreatic secretions are first produced by exocrine acinar cells, which are highly specialized cells that produce digestive enzymes including the zymogen proteases trypsinogen and chymotrypsinogen, pancreatic lipase, nucleases, and amylase.[69] These enzymes are secreted into the intra-pancreatic ductal system that drains into the main pancreatic duct. The main pancreatic duct then merges with the common bile duct and drains into the duodenum at the hepatopancreatic ampulla. The pancreatic ductal system is lined by cholangiocytes that participate in the modification of acinar cell secretions, similarly to bile duct cells that modify bile.[50] Pancreatic ductal cells, in particular, are responsible for secreting a bicarbonate rich solution that is essential for neutralizing the acidic chyme that enters the intestine from the stomach.[69] Pancreatic ducts also secrete mucins.[50] It is thought that many of the same molecular mechanisms involved in bile duct fluid and bicarbonate secretion are conserved in pancreatic ducts. Indeed, many of the same transporters in bile ducts have been shown to be expressed in pancreatic ducts, such as CFTR.[50,69] Further, pancreatic duct cells have been shown to

respond similarly to SCT, VIP, ACh, and bombesin as bile duct cells.[50,69] Interestingly, several groups have observed that rare endocrine hormone producing cells may reside in the pancreatic ductal compartment and could act to directly secrete hormones such as insulin, glucagon, and SST into the pancreatic exocrine secretions.[70–72]

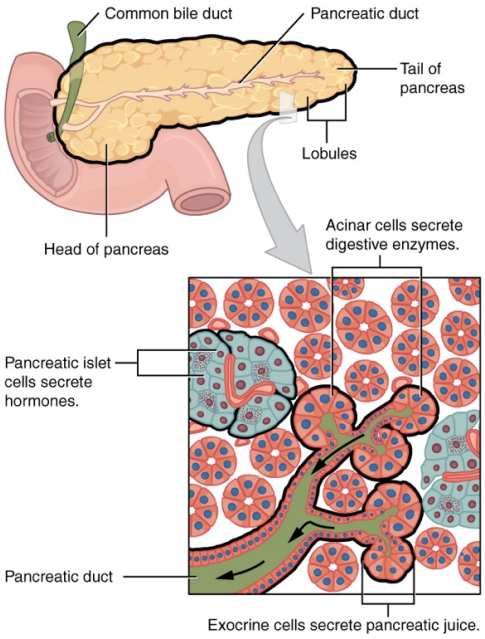


Figure 1.10 Diagram depicting the anatomy and microstructure of the pancreatic endocrine and exocrine system.

Image reproduced with permission from Anatomy and Physiology by OpenStax, Chapter 23.6 Accessory Organs in Digestion: The Liver, Pancreas, and Gallbladder

1.3.4 Embryology of the biliary and pancreatic duct systems

Developmentally, the liver and biliary tree derives from the ventral foregut endoderm which gives rise to the hepatic diverticulum.[73–76] However, the intrahepatic, extrahepatic, and pancreatic ductal compartments arise from separate progenitor populations which derive from the foregut endoderm. Biliary precursor cells, that will ultimately become intrahepatic cholangiocytes, form around the 8th week of digestion in humans.[73] They are derived from bipotential hepatoblasts in the liver bud that forms from the cranial portion of the hepatic diverticulum of the foregut endoderm (**Figure 1.11**).[73–77]

These bipotential hepatoblasts give rise to both hepatocytes and cholangiocytes (**Figure 1.12**).[74] Intrahepatic bile duct precursors form from the ductal plate.[73–75] The ductal plate is a ring-like structure formed from hepatoblasts adjacent to the developing portal vein mesenchyme that are induced to become cholangiocyte precursors. Ultimately, the ductal plate gives rise to tubular primitive intrahepatic bile duct structures that will become fully mature after birth.[75]

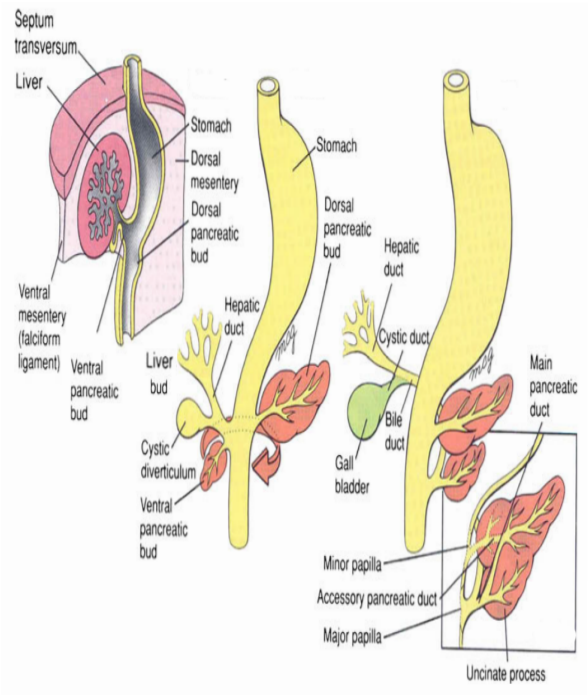


Figure 1.11 Diagram depicting the embryological development of the intrahepatic and extrahepatic biliary tree and the pancreas.

Reproduced with permission from Vakili and Pomfret (2008).[77]

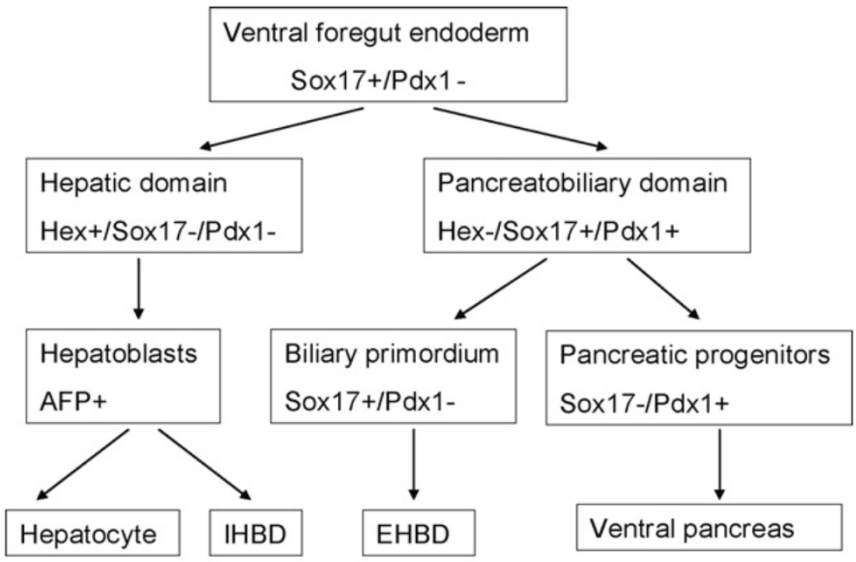


Figure 1.12 Developmental origin of intrahepatic and extrahepatic bile ducts, hepatocytes, and pancreas.

Diagram depicting the developmental cellular origin and key markers expressed by these progenitors for the intra- and extrahepatic bile ducts, hepatocytes, and ventral pancreas. Reproduced with permission from Zong and Stranger (2011).[74]

Several signalling pathways have been found to be important for the differentiation and development of intrahepatic cholangiocytes. Many of these signals are derived from the portal mesenchyme, which lies adjacent to the ductal plate (**Figure 1.13**).[73] One of the most important of these pathways is the Notch signalling pathway. It has been shown that Notch signalling is a master regulator of biliary specification factors HES1, SOX9, HNF1B, and HNF6, whose expression is important in committing hepatoblasts to a biliary fate. Periportal mesenchymal cells and developing cholangiocytes themselves express Jagged 1, the ligand for Notch receptors. In humans, mutations in either Jagged 1 or Notch 2 have been shown to cause congenital abnormalities in intrahepatic bile ducts, leading to bile duct paucity, a syndrome known as Alagille Syndrome.[73–76,78] Recent reports have also demonstrated the importance of both the Hippo and WNT signalling pathways in bile duct development. Not pictured in **Figure 1.13**, the Hippo downstream regulators, YAP/TAZ act to directly increase TGF-ß expression to commit hepatoblasts to a cholangiocyte fate.[79] Additionally, the tight regulation of WNT signalling and ß-Catenin levels are crucial for normal bile duct development and morphogenesis.[80]

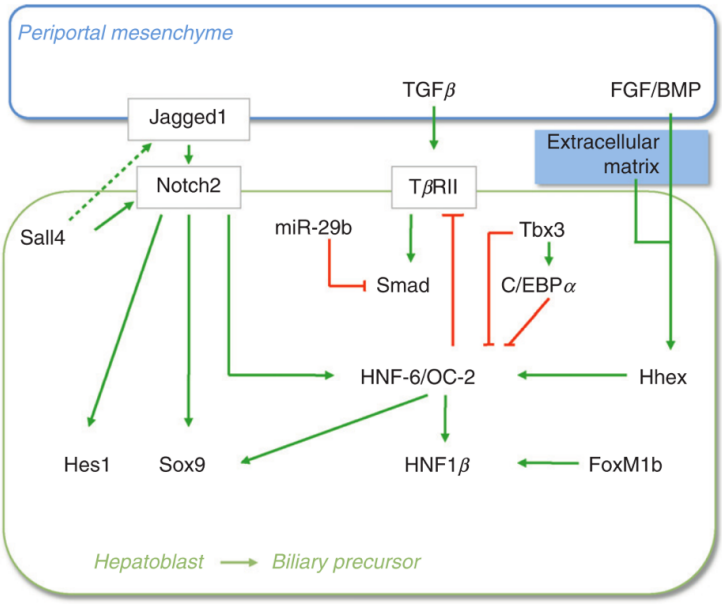


Figure 1.13 Molecular mechanisms controlling hepatoblast commitment to a biliary precursor fate.

Diagram depicting important molecular mechanisms controlling biliary commitment in developing liver. Image reproduced with permission from Lemaigre (2010).[73] Note that this figure is an overview and other important pathways such as Hippo and WNT are not shown here.

The development of the extrahepatic bile duct compartment has not been as well-studied as the intrahepatic compartment.[73–77] Little is known on the molecular mechanisms regulating extrahepatic development. However, several factors are common to both intra- and extrahepatic cholangiocyte development. The extrahepatic bile ducts and gallbladder arise before the intrahepatic biliary system. In the mouse, the extrahepatic and intrahepatic compartments are physically separate and anastomose later in development. However, in humans, the extrahepatic bile duct compartment is always in contact with the liver. The extrahepatic biliary tree and ventral pancreatic buds are derived from progenitor cells in the caudal portion of the hepatic diverticulum. These progenitor cells express both SOX17 and PDX1, whereas intrahepatic precursors do not express these markers.[73–76,78] These cells will eventually become specified towards either a biliary or pancreatic fate and SOX17 seems to control this.[78] It has been shown that SOX17 is necessary for the biliary specification of these cells. In the absence of SOX17, gallbladder agenesis occurs and common bile duct tissue is replaced with ectopic pancreatic tissue.[78] Likewise, overexpression of SOX17 leads to the ectopic development of biliary tissue in the pancreas, stomach, and duodenum.

Similar to intrahepatic bile ducts, there is evidence for the importance of HES1, HNF1B, and HNF6 in extrahepatic biliary specification.[73–76,78] In the absence of HES1, intrahepatic bile ducts do not develop normally and pancreas tissue replaces extrahepatic ductal tissues.[73–76,78]

The pancreas and pancreatic ductal systems are formed from a combination of the ventral pancreatic bud (derived from the hepatic diverticulum) and the dorsal pancreatic bud (derived directly from the foregut endoderm).[81,82] The main pancreatic duct is formed through fusion of the ventral pancreatic duct with the distal portion of the dorsal pancreatic duct and arises from PDX1+/SOX17- progenitor cells.[74,82] The proximal portion of the dorsal pancreatic duct often regresses and forms an accessory pancreatic duct which drains into the duodenum through the minor papilla. It does not anastomose with the common bile duct as the main pancreatic duct does.[81] Interestingly, this suggests that portions of the pancreatic ductal system may have a closer embryologic origin to the duodenum than to the cells arising from the hepatic diverticulum which give rise to the ventral pancreas, gallbladder, cystic duct, common bile duct and common hepatic duct.[82]

1.3.5 Diseases of the biliary and pancreatic duct systems

The human biliary tree is impacted by a range of disease states and many of these diseases show regionalization. These biliary disorders are collectively termed cholangiopathies and are responsible for 16% of all liver transplants and 50% of paediatric liver transplants.[83] When bile flow out of the liver is disturbed in these disease states, it leads to hepatotoxicity and ultimately can cause liver fibrosis and ESLD necessitating transplant. Therapeutic treatment options for most of these diseases are limited, as is the understanding of the underlying disease pathogenesis.[83] Several cholangiopathies preferentially target small intrahepatic bile ducts such as Primary Biliary Cirrhosis, an autoimmune inflammatory disorder. While others, such as Primary Sclerosing Cholangitis, Biliary Atresia, and non-anastomotic biliary strictures, target large intrahepatic ducts and/or extrahepatic bile ducts early in disease.[84,85] Cholangiocarcinoma of the intrahepatic bile ducts, extrahepatic bile ducts, and gallbladder represent unique entities and show phenotypical heterogeneity. It is unclear why these diseases target particular regions of the biliary tree.

However, it does suggest key differences may exist within these tissues. The regionalization of certain cholangiopathies is summarized in **Figure 1.14**.

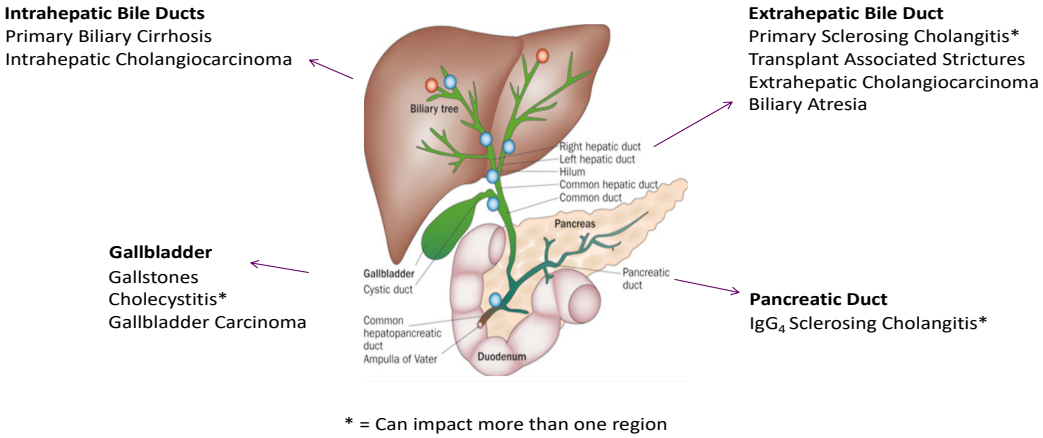


Figure 1.14 Regionalization of human cholangiopathies.

Overview of the regionalization and target areas of human cholangiopathies. This is a simplified view and represents the usual clinical presentation of the disease. However, several of these diseases can impact more than a single region of the biliary tree or have rarer subtypes of disease that may impact more regions than presented here. Image modified with permission from Cardinale et al (2012).[52]

1.4 Organoids

Organoid is a broad term, but generally is defined as any *in vitro* culture system which utilizes a three dimensional extracellular matrix in order to support the growth of self-organizing cellular structures.[86,87] Organoids can be derived from a broad range of cell types but are often derived from stem cells including cells differentiated from iPS/ES cells or adSCs. However, organoids do not have to be derived from stem cells and can be derived from mature cell types and even tumour cells.[88] Generally, organoid cultures share many features of the tissue types from which they were derived including tissue specific marker expression and functionality. Most organoid culture systems contain only one cell lineage, most often epithelial cells.[86,87] However, co-cultured organoids containing both epithelial and mesenchymal lineages have also been described.[86,87,89]

1.4.1 Adult stem cell derived organoids

Adult stem cell derived organoids are thought to be established from tissue resident adSCs.[16,87] The culture conditions used to grow adSC organoids attempt to recreate the *in vivo* stem cell niche. Alternatively, conditions may act to mimic states of tissue injury or inflammation which activate adSCs.[87] Remarkably, most adSC organoid systems rely on a well-conserved set of factors, with only minor differences between tissue types and even between species. As previously mentioned, one of the crucial factors in the ability to propagate adSCs *in vitro* as organoids was the discovery of the RSPO/LGR5/WNT axis.[25,87] adSC organoid systems rely heavily on a source of WNT and almost always include RSPO and often also a canonical WNT pathway agonist such as WNT3A or CHIR 99021 (a small molecule inhibitor of one of the destruction complex members, GSK3β) in the media. Additionally, it has been shown that modifying other signalling pathways is often integral to the establishment of organoid cultures. These include:

(1) Activation of the Epidermal Growth Factor (EGF) pathway
(2) TGF-ß inhibition (often accomplished using small molecule inhibitors such as A 83-01)
(3) Bone Morphogenic Protein (BMP) pathway inhibition (often accomplished by using the inhibitor, Noggin).
(4) An extracellular matrix (ECM) source, most commonly Matrigel. Matrigel is a semi-solid gel that is derived from murine Engelbreth-Holm-Swarm sarcoma tumour cell lines. At 4°C it is liquid but at 37°C it solidifies. This makes it an ideal ECM matrix source for embedding cells in three-dimensional (3D) *in vitro* culture systems.[90]

One of the first organoid systems to be described was for culturing small intestinal, LGR5+ adSCs.[21] This culture system was developed by Hans Clevers and colleagues and revolutionized the field of organoid research. Key to this discovery was decades of work invested in understanding the stem cell niche within the intestine. The intestine is a highly regenerative tissue with the human intestinal epithelium turning over completely in 2-3 weeks.[67] The intestinal epithelium consists of a highly organized structure, with villus and crypt domains that are organized to provide the maximal surface area for digestion (**Figure 1.15**).[67]

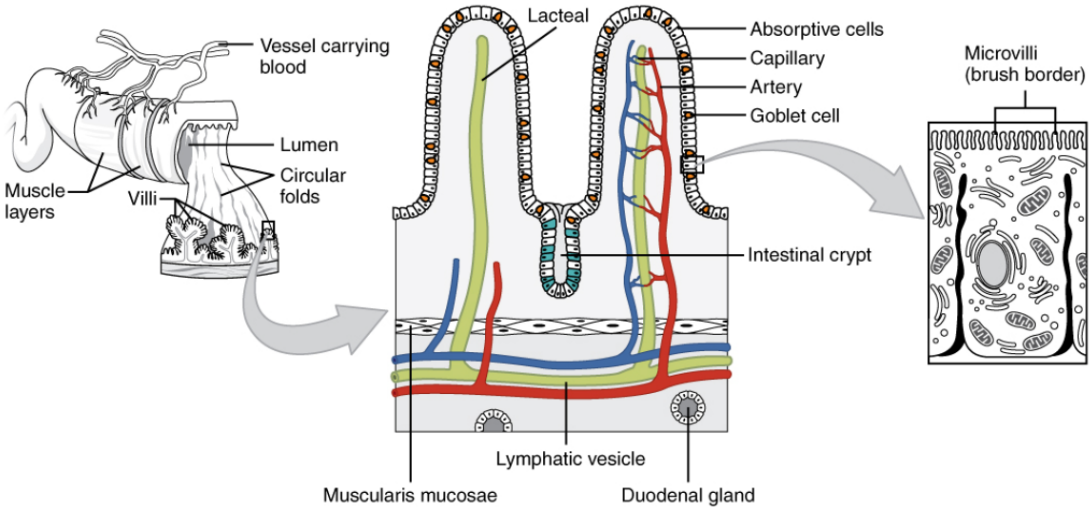


Figure 1.15 Anatomy and microstructure of the intestine.

Diagram reproduced with permission from Anatomy and Physiology by OpenStax, Chapter 23.5 The Digestive System: The Small and Large Intestines.[67]

The intestinal epithelium contains a heterogenous population of cell types including absorptive enterocytes, mucin producing goblet cells, paneth cells which secrete lysozyme and can act as phagocytes, and enteroendocrine hormone producing cells (i.e. S-cells which secrete Secretin or I-cells which secrete Cholecystokinin (CCK)).

It was discovered through lineage tracing studies that a LGR5+ stem cell population resides within the crypts of the intestine.[21,91] These stem cells are called crypt based columnar cells (CBCs) and are capable of giving rise to all of the differentiated cell types of the intestine (**Figure 1.16**). As LGR5+ cells divide to produce daughter cells, they migrate up the crypt towards the villi tip. Signalling gradients are produced by mesenchymal and other niche cells that lead to progressive differentiation of these cells as they migrate up the villi. These intermediately differentiated, proliferative cells are called transit-amplifying (TA) cells. This understanding of intestinal regeneration, as well as the discovery of RSPO/LGR5, led to the discovery that intestinal organoids could be cultured *in vitro*.

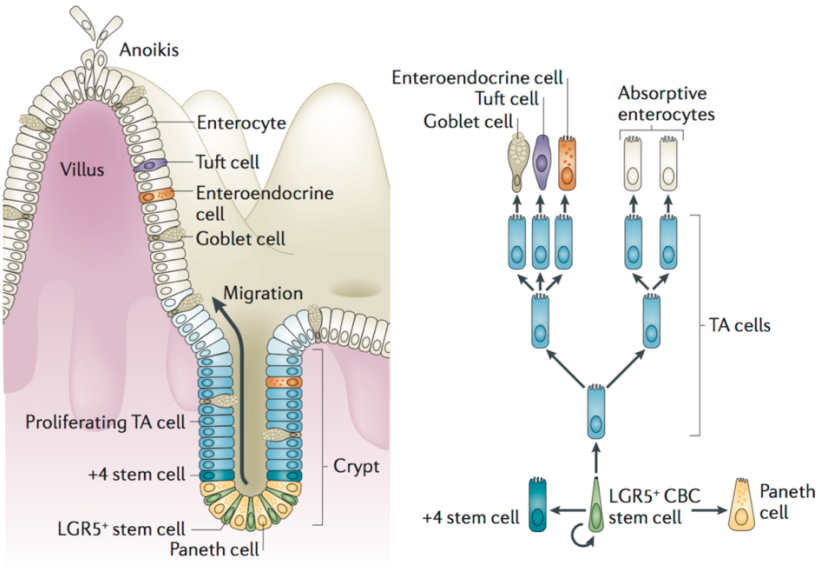


Figure 1.16 Adult stem cell niche of the intestine

Diagram depicting the adult stem cell niche of the intestine and the process by which LGR5+ stem cells divide to give rise to proliferative transit-amplifying (TA) cells which ultimately differentiate to give rise to the mature cell types of the intestine. Diagram reproduced with permission from Barker (2014).[92]

These organoids, which could be grown from single LGR5+ cells embedded in matrigel, remarkably self-organize into "mini-guts", with budding structures reminiscent of the crypt/villus domain of the adult intestine. Furthermore, *in vitro* intestinal organoids not only contain LGR5+ stem cells but also give rise to all the differentiated cell lineages of the intestine *in vitro.* The composition and percentage of these differentiated cells can be altered and NOTCH and WNT signalling pathways appear crucial for this.[93] Following the discovery of intestinal organoids, it was found that many other adult tissues from endoderm, mesoderm, and ectodermal tissues are capable of forming *in vitro* organoids and that many of these tissues also express the intestinal CBC cell marker LGR5. adSC organoids have been derived from tissues including the mammary glands, prostate, lung, salivary gland, stomach, large intestine, and oesophagus.[20,24,28,87,88] Recently it has also been shown in both mice and humans that organoids, potentially containing adSCs, can be derived from intrahepatic bile ducts (often termed Liver Organoids), intrapancreatic ducts (often termed Pancreas Organoids), and even the gallbladder.[94–98] These biliary organoids will be discussed below.

1.4.2 Biliary stem/progenitor cells and organoids

1.4.2.1 Liver progenitor cells/oval cells

Within the biliary tree, several groups have reported the existence of candidate stem/progenitor cells. Most of this work has historically focused on the intrahepatic bile duct compartment, in particular liver progenitor/oval cells. As mentioned above, in mild, acute injuries to the liver, such as partial hepatectomy, mature hepatocytes and cholangiocytes re-enter the cell cycle and divide to replace the lost cell mass. Whereas in chronic or severe acute liver injuries, it has been demonstrated that a population of bipotent progenitor cells of biliary origin may be responsible for regenerating lost hepatocyte and cholangiocyte mass.[36] These cells, called liver progenitor cells (LPCs), have been identified in both mice and humans and are thought to be capable of differentiating towards either cholangiocyte or hepatocyte fates.[43,44,99–102] However, the role that these cells play in regeneration and repair of the liver has become controversial in the field, with several recent studies indicating that biliary cells do not differentiate into hepatocytes in liver injury models to any appreciable degree.[103] A recent report by Forbes and colleagues may have definitively answered this debate by showing when murine hepatocytes are induced to become senescent and prevented from dividing, as is found in many human disease states, cells from the biliary compartment are capable of replacing these lost hepatocytes.[38]

The first liver progenitor cell population to be identified, termed oval cells, was discovered in rats in 1956 by Farber.[42] These cells were shown to appear after acute liver injury in periportal regions and had oval shaped nuclei.[42] Further research by multiple groups including Evarts et al. 1989 and Lazaro et al. 1998, suggested that these cells are bipotent, with the ability to give rise to hepatocytes and cholangiocytes *in vivo* and *in vitro.* They are thought to arise from a population of cells within intrahepatic bile ducts.[43,99] Similar cells have since been identified in human liver samples from a multitude of disease states as well as healthy livers.[44,97,100–102] Despite the discovery of oval cells in rodents and hepatic progenitor cells in humans, research on these cells was historically hindered due to the lack of specific molecular markers necessary for identifying these cells and allowing for their isolation.[43] However, recent research has begun to identify several possible markers of LPCs including: SOX9, FOXL1, SPP1, PROM1, CD24 and most recently LGR5.[45,97,98,104–106] Furthermore, recent research has also begun to identify important elements of the niche occupied by hepatic progenitor

cells and a more extensive understanding of these cells has been emerging. These cells are thought to reside in the canals of Hering, which is a boundary zone between intrahepatic bile ducts and hepatic cords surrounded by a laminin sheath.[37,107,108]

1.4.2.2 Intrahepatic biliary organoids

Within the biliary compartment, several organoid culture systems have been described. Huch et al. (2013) used lineage tracing to discover a population of cells in the intrahepatic biliary compartment of damaged livers of mice which may represent an oval cell/liver progenitor population.[98] In the mouse liver, these cells appeared in periportal regions following CCl_4 induced liver damage but were absent in normal livers. This population of cells was shown to express LGR5 and could be cultured *in vitro* in 3D organoid conditions for over 12 months. The organoids were cultured in media promoting WNT signalling, which contained RSPO and EGF, among other factors. These cells were subjected to differentiation protocols in an attempt to generate functional hepatocytes *in vitro* and were found to be capable of expressing low-levels of hepatocyte markers.[86] These differentiated cells were then transplanted into a mouse $FAH^{-/-}$ model to further test their differentiation ability. Engraftment rates however were low (0.1% to 1%) and phenotypic rescue was not demonstrated, although, survival in animals where engraftment took place was prolonged and transplanted cells appeared to take on hepatocyte-like fates *in vivo*.

Following this study, Huch and colleagues extended their work to human liver tissue in 2015.[97] From human livers, they were able to isolate intrahepatic biliary ducts and culture them as organoids in conditions promoting WNT signalling. Like their mouse counterparts, these cells expressed LGR5 and demonstrated long-term proliferative ability. These intrahepatic organoids also showed the capability of expressing hepatocyte specific markers under differentiation conditions. However, the expression of these markers was lower than adult liver tissue. The human cells were capable of engrafting in a mouse liver injury model *in vivo* and produced detectable levels of human albumin for 120 days post-transplantation.

Taken together, these reports strongly suggest that a stem/progenitor population exists in the intrahepatic bile duct compartment of both mouse and human livers. These cells can be cultured as 3D organoids and express markers of adSCs. However, it is important to note that it is still difficult to definitively conclude whether this population of cells derives

solely from intrahepatic bile ducts. In the mouse, bipotent periportal cells deriving from hepatocytes, and not cholangiocytes, have been reported to also participate in liver regeneration after injury.[109]

1.4.2.3 Extrahepatic biliary organoids

Despite the decades of work seeking to identify an intrahepatic biliary stem/progenitor cell, only a handful of studies have examined the extrahepatic compartment.[53,54,110] Several groups have suggested the existence of a stem/progenitor cell population more distally in the extrahepatic biliary tree.[53,54,110–117] These cells, similarly to intrahepatic cells detailed above, are thought to express markers of adSCs, such as LGR5, and may even be capable of differentiating towards other endodermal lineages, such as hepatocytes and pancreatic endocrine cells, *in vitro.*[53,54,110–117] Additionally, several reports have shown that extrahepatic and pancreatic biliary cells can be cultured as organoids.

Within the pancreas, it has been reported that intrapancreatic ducts can give rise to organoids. Huch et al (2013) described the derivation of intrapancreatic organoids from mice.[96] Similarly to their work in the liver, Huch and colleagues demonstrated that, following injury to the pancreatic ductal system using pancreatic duct ligation, a population of LGR5+ cells emerge from the intrapancreatic ducts. These LGR5+ cells could be subsequently isolated and cultured as organoids and relied on a RSPO based culture media. The organoids expressed markers of ductal cell lineages and markers commonly expressed in intestinal stem cells. These organoids could be expanded long-term, *in vitro,* and when aggregated with embryonic pancreatic cells and transplanted under the kidney capsule in mice, were shown to generate predominately ductal structures but also appeared capable of generating a low-level of endocrine hormone producing cell types. This work was subsequently extended to include the derivation of human intrapancreatic organoids and also pancreatic adenocarcinoma organoids.[95,118] Together, these reports suggest that a progenitor population may exist in the intrapancreatic ducts. However, the differentiation capacity of these cells is still unclear. Common to all of these reports, is that these cells, if they do differentiate beyond a ductal fate, do so at an extremely low efficiency. It also remains unclear how functional these cells truly are.[95,119] Interestingly, a report by Dorrell et al (2014) compared the organoid initiating cells from murine intrahepatic and intrapancreatic ducts. They found that the two cell types were very similar in surface marker expression and that intrapancreatic organoids were also

capable of engrafting in the livers of $FAH^{-/-}$ mice, suggesting similarities between these organoid systems may exist.[120]

Several groups have reported the possibility of culturing extrahepatic ductal organoids from the common bile duct and gallbladder. Lugli et al (2016) reported on the derivation of gallbladder derived organoids from both mouse and human tissues. These gallbladder organoids, similarly to intrahepatic and intrapancreatic organoids, relied on culture conditions containing RSPO. The murine gallbladder organoids were subjected to a differentiation protocol and were shown to upregulate expression of the hepatocyte marker Albumin in a small subset of cells. However, when the organoids were transplanted *in vivo,* under the liver capsule, they exclusively generated ductal lineages. Further, no differentiation potential was assessed for the human gallbladder organoids.

Additional reports by Manohar et al (2011 and 2015) identified a clonogenic population of potential progenitor cells in the adult mouse gallbladder and fetal human gallbladder.[115,117] These cells were cultured in two-dimensional (2D) cultures on fibroblast feeders. When transferred into a matrigel 3D-growth assay containing RSPO, they formed cystic and tubular organoid structures. Furthermore, these clonogenic gallbladder cells were shown to be distinct from intrahepatic biliary cells and expressed different surface markers as well as showed transcriptional differences. Additionally, work by Lola Reid and colleagues has suggested that a population of multipotent, endodermal progenitor cells may exist in the human extrahepatic bile duct and gallbladder.[111,113] These cells have been derived from EPCAM positive cells in adult gallbladder and common bile duct and may have the capacity to differentiate towards hepatic or pancreatic endocrine lineages, albeit at a low level. It is important to note that these cells display a limited proliferation capacity *in vitro* and until recently could not be serially passaged.[121] Further, the cells are reported to express markers of pluripotency such as OCT4, however, this finding has never been confirmed by other studies and is surprising given the absence of these pluripotent markers in adult tissues.

Lastly, our own group recently reported on an *in vitro* organoid system capable of expanding human common bile duct and gallbladder epithelial cells.[122] These organoids were grown in a 3D matrigel system with RSPO1, EGF, and DKK-1 (a canonical WNT pathway inhibitor). These organoids, called Extrahepatic Cholangiocyte Organoids (ECOs), displayed long-term proliferative ability without expressing markers associated with adSCs and instead appeared phenotypically to represent a functional, differentiated cholangiocyte population. Further supporting this was that ECOs were capable of

colonizing synthetic bio-scaffolds that were subsequently used to reconstruct the murine gallbladder and common bile duct. These organoid cells remained integrated in the scaffolds *in vivo* and maintained a mature cholangiocyte phenotype.[122]

Ultimately, as can be seen from the summary above, the literature relating to biliary organoids/progenitor cells is mixed. Some reports suggest the existence of progenitor cells capable of differentiation beyond a biliary fate throughout the biliary tree, while others suggest a more limited differentiation capacity for these cells, especially in the extrahepatic biliary tree. Additionally, the relationship between these numerous systems and the organoids derived from these different tissues is still largely unknown. Therefore, the goal of this dissertation was to address some of these key questions and is further detailed below.

1.5 Dissertation Objectives

The overarching goal of this dissertation was to explore what regionalization may exist in the human biliary tree. We sought to not only understand whether differences exist between primary tissue regions *in vivo* but also what differences may exist in cells cultured as *in vitro* organoids from these tissues.

To accomplish this, I first developed an *in vitro* organoid culture system capable of expanding cells from both intrahepatic and extrahepatic/pancreatic biliary tissues. Chapters 3 and 4 aim to address the following questions using these organoid systems and primary tissues:

(1) Can a population of cells with an adult stem/progenitor cell phenotype be isolated from the human common bile duct, gallbladder, and pancreatic duct (referred to collectively as extrahepatic tissues/cells)? Can these extrahepatic cells be expanded long-term *in vitro* in an organoid culture system promoting WNT signalling*?*
(2) How do the cells isolated from different regions of the extrahepatic biliary tree compare to their tissue of origin, to each other, and to intrahepatic bile duct organoids?
(3) Do different anatomical regions of the biliary tree demonstrate unique transcriptional profiles?

Finally, Chapter 5 aims to describe the differentiation capacity of these organoid systems and in particular aims at addressing the following question:

(4) Do these extrahepatic cells represent a stem/progenitor population with bi- or multipotentiality? If so, what is the differentiation capacity of these cells?

Overall, the work of this dissertation seeks to advance the understanding of biliary physiology and the understanding of how different regions of the human biliary tree may differ *in vivo* and in *in vitro* organoid systems. Secondly, it aims to assess whether human extrahepatic biliary tissues are capable of giving rise to organoid cultures with a progenitor cell phenotype and whether they are capable of differentiation towards a hepatocyte fate, as has been suggested with intrahepatic tissues.

2 MATERIALS AND METHODS

2.1 Statement of Source

The methods presented here are largely based on experiments presented in the following first author manuscript written by the author of this dissertation. Therefore, some parts of these methods have been taken *verbatim* or with only minor changes from this source.

Rimland CA, Tilson S, Morell C, Tomaz R, Lu WY, Adams S, Georgakopoulos N, Otaizo-Carrasquero F, Myers T, Sun HW, Gieseck RL, Sampaziotis F, Tysoe O, Wesley B, Oniscu GC, Hannan NRF, Forbes S, Saeb-Parsy K, Wynn TA, Vallier L. (2018) Regional differences in human biliary tissues and corresponding *in vitro* derived organoids. Manuscript in preparation.

2.2 Human Tissue Material

Human GBD, CBD, PancD, and liver samples were obtained from organ donors at either Cambridge University Hospitals or the Royal Infirmary of Edinburgh in the United Kingdom. CBD, PancD, and liver samples were obtained when a liver or pancreas was deemed unsuitable for transplant, whereas GBD samples were obtained even when a liver was suitable for transplant, as the GBD is not routinely transplanted. Informed consent for the use of human tissues for research purposes was obtained from each donor's next of kin and protocols were reviewed and approved by the ethics committees at both hospitals and consent recorded according to the NHSBT consent regulations.

Tissue samples were stored on ice in University of Wisconsin Solution until processed for cell isolation as described below. Donor demographics and corresponding cell lines derived from each tissue sample are listed in **Table 2.1**.

Tissue	Donor Demographics	Tissue Sample / Cell Line	RNA-Seq T=tissue O=organoids	Internal ID
Gallbladder	21, Male	Tissue and Cell Line (-A8301/FSK)	Not sequenced	GBD1
Gallbladder	33, Female	Tissue and Cell Line (-A8301/FSK)	Not sequenced	GBD2
Gallbladder	39, Male	Tissue and Cell Line	Not sequenced	GBD3
Gallbladder	21, Female	Tissue and Cell Line	T	GBD4
Gallbladder	55, Female	Tissue and Cell Line	Not sequenced	GBD5
Gallbladder	44, Male	Tissue and Cell Line	T	GBD6
Gallbladder	63, Male	Tissue and Cell Line	T/O	GBD8
Gallbladder	61, Female	Tissue and Cell Line	T/O	GBD10
Gallbladder	67, Male	Tissue and Cell Line	T/O	GBD11
Gallbladder	67, Male	Tissue and Cell Line	Not sequenced	GBD12
Gallbladder	68, Female	Cell Line only	Not sequenced	GBD13
Common Bile Duct	50, Female	Cell Line only (-A8301/FSK)	Not Sequenced	CBD2
Common Bile Duct	63, Male	Tissue and Cell Line	T/O	CBD4
Common Bile Duct	20, Male	Tissue and Cell Line	T/O	CBD5 (PD4)
Common Bile Duct	67, Male	Cell Line only	O	CBD6
Common Bile Duct	58, Female	Tissue only	T	CBD8
Common Bile Duct	43, Female	Tissue Only	T	PD1
Common Bile Duct	48, Female	Tissue and Cell Line	T/O	PD2
Common Bile Duct	36, Male	Tissue and Cell Line	T/O	PD3
Common Bile Duct	48, Female	Tissue and Cell Line	T/O	PD5
Common Bile Duct	61, Female	Cell Line only	Not sequenced	PD6
Pancreatic Duct	42, Male	Tissue and Cell Line	T/O	PancD1
Pancreatic Duct	27, Female	Tissue and Cell Line	T/O	PancD2
Pancreatic Duct	37, Male	Cell Line only	O	PancD4
Pancreatic Duct	53 , Male	Tissue only	T	PancD5
Liver	50, Male	Tissue and Cell Line IHBD_ex, IHBD_Huch	O-IHBD_Huch O- IHBD_ex	IHBD1
Liver	30, Female	Tissue and Cell Line IHBD_ex	O-IHBD_ex	IHBD2
Liver	50, Female	Tissue and Cell Line IHBD_ex	O-IHBD_ex	IHBD3
Liver	67, Female	Cell Line only IHBD_ex, IHBD_Huch	O-IHBD_Huch	IHBD4
Liver	76, Male	Cell Line only IHBD_ex, IHBD_Huch	O-IHBD_Huch O-IHBD_ex	IHBD5

Table 2.1 Donor Demographics

Details of donor tissue samples used for organoid cell line generation or for collection of primary tissue RNA.

2.3 Extrahepatic epithelial cell isolation and organoid culture

GBD samples were received as either whole intact GBDs or small 2 cm^2 tissue segments. Whole GBDs were first punctured with a scalpel and drained of bile. The neck of the GBD was then dissected away and discarded. A longitudinal incision was used to open the GBD and expose the mucosal surface. No further processing was required for the GBD segment samples. CBD and PancD samples were received intact. To expose the bile duct lumen the wall of the duct was incised with a scalpel. All samples were washed 3 times with ice cold Hanks Balanced Salt Solution (HBSS) without Ca^{2+}, Mg^{2+}, or Phenol Red. Tissues were then transferred to cold ADV/DMEM F12 media supplemented with 100 U/mL penicillin and 100 µg/mL streptomycin. The mucosa was then abraded with a scalpel to mechanically dissociate the cells (**Figure 2.1**).

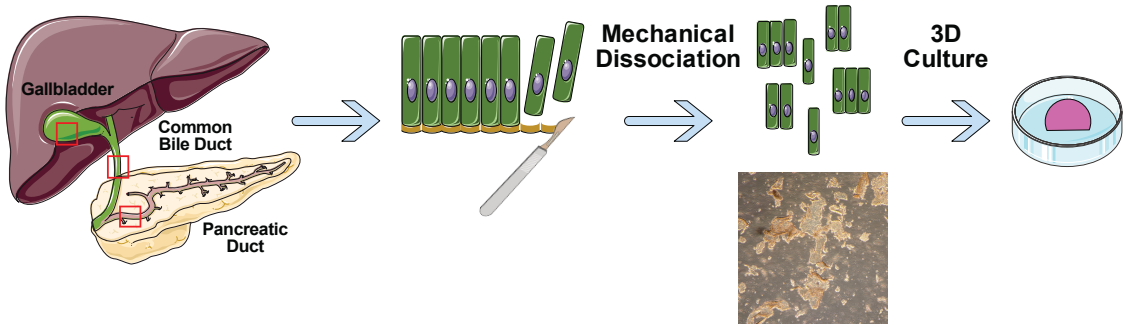


Figure 2.1 Extrahepatic biliary epithelial cell isolation procedure.

Diagram depicting the mechanical isolation method employed to obtain biliary epithelial cells and image showing the sheets of epithelial cells recovered.

Cell suspensions were centrifuged at 300 x g for 5 min, washed twice, and re-suspended in growth factor reduced matrigel. 50 µL droplets were plated per well into a 24-well plate. Matrigel was allowed to solidify for 5-10 min at 37°C before adding 500 µL of ADV/DMEM F12 media containing 1x N2-serum free supplement, 1x B27-serum free supplement, 2 mM L-glutamine, 100 U/mL penicillin, and 100 µg/mL streptomycin. The following growth factors were added: 20% R-Spondin conditioned media or 500 ng/mL recombinant human R-Spondin 1, 3 µM CHIR 99021, 100 ng/mL recombinant human noggin, 2.5 µM PGE_2, 100 ng/mL recombinant human EGF, 5 µM A 83-01, and 10 µM forskolin. For the first 48 h, 10 µM Y-27632 was added to the media. Media was changed

every 2-3 days. Within 24-48 hr, small cystic organoid structures could be observed. Organoids reached confluence after 5-10 days and were split using the procedure below. Cells could be readily freeze/thawed using Cell Banker 2 freezing media.

2.4 Intrahepatic bile duct isolation and organoid culture

Small pieces of human liver tissue (2x2 cm^2) were frozen in Cell Banker 2 and stored at -80°C until duct isolation was required. For duct isolation (**Figure 2.2**), frozen liver tissue was removed from -80°C and thawed in a 37°C water bath. The thawed tissue was then minced into small pieces using a scalpel blade in a 10cm^2 tissue culture plate. The tissue fragments were washed off the plate and placed into a 50mL conical tube with 20mL of digestion media. Digestion media consisted of DMEM High Glucose + Glutamax with 1% Fetal Calf Serum (FCS), 100 U/mL penicillin, 100 µg/mL streptomycin, 128.4 µg/mL of dispase II and 128.4 µg/mL of collagenase from *Clostridium histolyticum*. Liver tissue in digestion media was placed in a 37°C water bath. Every 20 min, for 4-5 hr, the tissue was disrupted with a serological pipette of decreasing volume (25mL, 10mL, 5mL, 2mL) and ultimately a p1000 pipette to release ducts from surrounding liver tissue. Every 1-2 hr the tissue pieces were allowed to settle to the bottom of the tube and the digestion medium was replaced with new media.

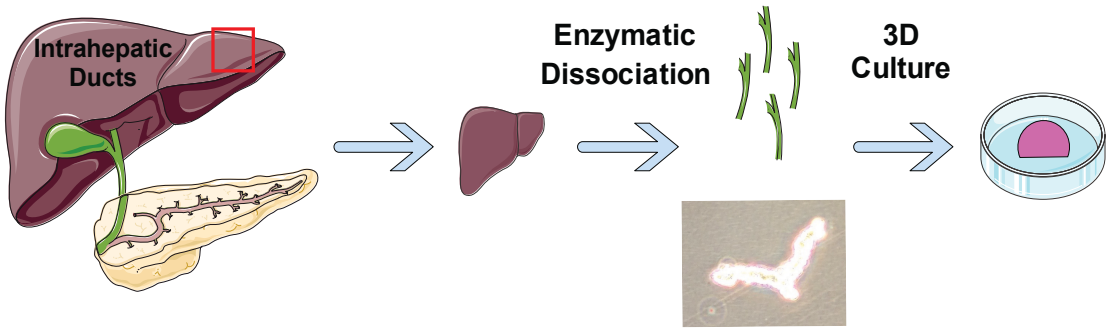


Figure 2.2 Intrahepatic bile duct isolation procedure

Diagram depicting the isolation procedure for intrahepatic bile ducts and image showing an isolated duct.

After 4-5 hr the suspension became clear and could easily pass through a p1000 pipette. Ducts were manually picked under a microscope using a p200 pipette and placed into a tube containing DMEM High Glucose + Glutamax media with 5% FCS. Ducts were

centrifuged for 5 min at 200 x g. Supernatant was aspirated and ducts were re-suspended in growth factor reduced matrigel. 50μL droplets of matrigel were plated in a 24-well culture plate and allowed to solidify at 37°C. Ducts were overlaid with 500uL of one of two medias (1) extrahepatic culture media as described above with 10 μM Y-27632 or (2) intrahepatic isolation media for 3 days after which the media was changed to intrahepatic expansion media as described in Huch et al (2015) and Broutier et al (2016).[97,123] See **Table 2.2** for all reagents and manufacturers used in cell culture.

2.5 Organoid passaging

Extrahepatic organoid cultures were split every 5-10 days at a ratio of 1:2-1:4. Culture media was removed and cells were incubated in cell recovery solution for 30 min at 4°C. Extrahepatic organoids were washed 2 times with ADV/DMEM F12 media and pelleted at 300 x g for 4 min. Extrahepatic organoids were mechanically dissociated into small clumps and re-suspended in fresh matrigel and re-plated as above. For the first 2 days after splitting, media was supplemented with 10 μM Y-27632.

Intrahepatic organoids were subjected to a similar passaging procedure, with one exception. Intrahepatic organoids were not passaged using cell recovery solution and were instead mechanically dissociated in cold ADV/DMEM F12 media, disrupted with a p200 pipette, washed once and pelleted at 300 x g for 3 min. All other steps remained the same as detailed above for extrahepatic organoids.

Basal Organoid Medium = ADF+	**Final Concentration**	**Supplier (catalog number)**
Advanced DMEM/F12 (ADF)	N/A	Thermo Fisher (12634010)
N2 Supplement 100X	1 X	Thermo Fisher (17502048)
B27™ Supplement 50X, serum-free	1 X	Thermo Fisher (17504044)
L-Glutamine	2 mM	Thermo Fisher (25030081)
Penicillin/Streptomycin	100 U/mL / 100 µg/mL	Thermo Fisher (15140122)
Extrahepatic Organoid Expansion Media (EHBD organoids or IHBD_ex organoids)	**Final Concentration**	**Supplier (catalog number)**
ADF+	See above	See above
RSPO-conditioned media or Recombinant Human R-Spondin 1	20% (vol/vol) 500 ng/mL	Cambridge Stem Cell Institute (homemade) R&D (4645-RS)
CHIR 99021	3 µM	Tocris (4423)
Recombinant Human Noggin Fc Chimera	100 ng/mL	R&D (3344-NG)
Prostaglandin E_2	2.5 µM	R&D (2296)
Recombinant Human EGF, carrier free	100 ng/mL	R&D (236-EG)
A 83-01	5 µM	Tocris (2939)
Forskolin (FSK)	10 µM	Sigma (F6886)
Rock Inhibitor Y-27632 (added for first two days after isolation or splitting)	10 µM	Selleck Chem (S1049)
Intrahepatic Organoid Isolation Media (IHBD_Huch organoids for first 3 days after isolation)	**Final Concentration**	**Supplier (catalog number)**
ADF+	See above	See above
RSPO-conditioned media	10% vol/vol	Cambridge Stem Cell Institute (homemade)
Wnt-conditioned media	30% vol/vol	Cambridge Stem Cell Institute (homemade)
Recombinant Human Noggin Fc Chimera	25 ng/mL	R&D (3344-NG)
N-acetylcysteine	1.25 mM	Sigma (A9165)
Nicotinamide	10 mM	Sigma (N0636)
Human Gastrin I	10 nM	Sigma (G9145) or R&D (3006/1)
Recombinant Human EGF, carrier free	50 ng/mL	R&D (236-EG)
Recombinant Human HGF	25 ng/mL	Peprotech (100-39)
Recombinant Human FGF-10	100 ng/mL	Peprotech (100-26) or Autogen BioClear (ABC144)
A 83-01	5 µM	Tocris (2939)
Forskolin (FSK)	10 µM	Sigma (F6886)
Rock Inhibitor Y-27632	10 µM	Selleck Chem (S1049)

Table 2.2 Reagents used for cell culture

Intrahepatic Organoid Expansion Media (IHBD_Huch organoids)	**Final Concentration**	**Supplier (catalog number)**
ADF+	See above	See above
RSPO-conditioned media	10% vol/vol	Cambridge Stem Cell Institute (homemade)
N-acetylcysteine	1.25 mM	Sigma (A9165)
Nicotinamide	10 mM	Sigma (N0636)
Human Gastrin I	10 nM	Sigma (G9145) or R&D (3006/1)
Recombinant Human EGF, carrier free	50 ng/mL	R&D (236-EG)
Recombinant Human HGF	25 ng/mL	Peprotech (100-39)
Recombinant Human FGF-10	100 ng/mL	Peprotech (100-26) or Autogen BioClear (ABC144)
A 83-01	5 µM	Tocris (2939)
Forskolin (FSK)	10 µM	Sigma (F6886)
Rock Inhibitor Y-27632 (added only for first two days after splitting)	10 µM	Selleck Chem (S1049)
Huch et al (2015) Differentiation Media Phase I	**Final Concentration**	**Supplier (catalog number)**
Intrahepatic Organoid Expansion Media	See above	See above
Recombinant Human BMP-7	25ng/mL	Peprotech (120-03)
Huch et al (2015) Differentiation Media Phase II	**Final Concentration**	**Supplier (catalog number)**
ADF+	See above	See above
N-acetylcysteine	1.25 mM	Sigma (A9165)
Human Gastrin I	10 nM	Sigma (G9145) or R&D (3006/1)
Recombinant Human EGF, carrier free	50 ng/mL	R&D (236-EG)
Recombinant Human HGF	25 ng/mL	Peprotech (100-39)
A 83-01	0.5 µM	Tocris (2939)
Recombinant Human BMP-7	25ng/mL	Peprotech (120-03)
Dexamethasone	3 µM	Sigma (D4902)
DAPT	10 µM	Sigma (D5942)
Recombinant Human FGF19, carrier free	100 ng/mL	R&D (969-FG/CF)
Additional Reagents used for cell culture and liver tissue digestion	**Final Concentration**	**Supplier (catalog number)**
Cell Banker 2	N/A	Amsbio (11891)
Growth Factor Reduced Matrigel (with or without phenol red)	100%	Corning (354230) or (356231)
Cell Recovery Solution	N/A	Corning (354253)
TrpLE Express	N/A	Thermo Fisher (12604013)
Dispase II	128.4 µg/mL	Gibco (17105-041)
Collagenase from *Clostridium histolyticum*	128.4 µg/mL	Sigma (C9407)
DMEM High Glucose + Glutamax	N/A	Thermo Fisher (10566016)

Table 2.2 (continued) Reagents used for cell culture

2.6 Flow cytometry

Primary epithelial cell scrapings and/or organoid cultures were dissociated to single cells using TrypLE Express for 5-10 min at 37°C. Cells were washed once with 1% BSA to remove any remaining TrypLE and centrifuged at 400 x g for 3 min. Cells were then fixed with 4% Paraformaldehyde (PFA) for 20 min on ice. Cells were washed twice with 1% Bovine Serum Albumin (BSA) and pelleted at 400 x g for 3 min and re-suspended in 1% BSA and stored at 4°C until needed. Cells were blocked with 10% donkey serum/0.1% Triton-X for 30 min at room temperature. Primary antibodies, diluted in 1% donkey serum/0.1% Triton-X (antibody diluent), were applied at room temperature for 1 hr. Cells were washed 3 times with antibody diluent. Secondary antibodies, diluted in antibody diluent, were applied at room temperature for 1 hr. Cells were washed 3 times with antibody diluent and filtered through a 40 μm mesh filter prior to sorting. Gates were set using single stained, fluorescence minus one (FMO) controls and with cells stained with secondary antibodies alone. 10,000 or greater events were captured for each sample.

2.7 Immunocytochemistry

2.7.1 Whole mount organoids

Organoids in matrigel were grown in 48-well tissue culture plates (Corning) on 9 mm sterilized round glass coverslips (VWR). Culture media was carefully removed and organoids were washed for 5 min in Phosphate Buffered Saline (PBS). Organoids were then fixed in 4% PFA for 20 min at room temperature to maintain integrity of the matrigel. Organoids were washed 2 times with PBS to remove any remaining PFA. Organoid coverslips were blocked and permeabilized in 10% donkey serum/0.1% Triton-X for 3 hr. Primary antibodies, diluted in 1% donkey serum/0.1% Triton-X (antibody diluent), were applied overnight at 4°C. Organoid coverslips were washed 3 times at room temperature in the antibody diluent for 1 hr per wash. Secondary antibodies were applied overnight at 4°C and then washed 2 times for 1 hr each at room temperature. Nuclei were stained with Hoechst 33258 (1:10,000, Sigma) in PBS for 30 min at room temperature.

Organoid coverslips were washed once more with PBS for 1 hr. Coverslips were carefully removed using forceps and mounted on glass slides using Fluoromount G. A list of antibodies used can be found in **Table 2.3**.

Antibody	Species	Company/Product #	Dilution	Application
EPCAM	Mouse Monoclonal	R&D / MAB9601	1:100	IC IC-Fr FC
KRT7	Rabbit Monoclonal	AbCam / ab68459	1:100	IC IC-Fr FC
KRT7	Mouse Monoclonal	AbCam / ab9021	1:100	IC IC-Fr
KRT19	Mouse Monoclonal	AbCam / ab7754	1:100	IC IC-Fr FC
HNF4α	Rabbit Monoclonal	AbCam / ab92378	1:100	IC IC-Fr
Sox9	Rabbit Monoclonal	AbCam / ab185230	1:100	IC IC-Fr
β-Catenin	Goat Polyclonal	R&D / AF1329	1:100	IC IC-Fr
Ki67	Rabbit Monoclonal	AbCam / ab15580	1:100	IC IC-Fr
CDX2	Rabbit Monoclonal	AbCam / ab76541	1:100	IC IC-Fr
SOX17	Goat Polyclonal	R&D / AF1924	1:100	IC IC-Fr
Alexa Fluor 488	Donkey anti-goat IgG	Invitrogen / A11055	1:1000	IC IC-Fr
Alexa Fluor 488	Donkey anti-mouse IgG	Invitrogen / A10037	1:1000	IC IC-Fr FC
Alexa Fluor 488	Chicken anti-mouse IgG	Thermo / A21200	1:1000	IC IC-Fr
Alexa Fluor 488	Chicken anti-goat IgG	Thermo / A21467	1:1000	IC IC-Fr
Alexa Fluor 568	Donkey anti-goat IgG	Invitrogen / A11057	1:1000	IC IC-Fr
Alexa Fluor 568	Donkey anti-mouse IgG	Invitrogen / A10037	1:1000	IC IC-Fr
Alexa Fluor 568	Donkey anti-rabbit IgG	Invitrogen / A10042	1:1000	IC IC-Fr
Alexa Fluor 647	Donkey anti-rabbit IgG	Invitrogen / A31573	1:1000	FC
TRITC	Donkey anti-Goat IgG	Thermo / A16010	1:1000	IC IC-Fr
TRITC	Donkey anti-rabbit IgG	Thermo / A16040	1:1000	IC IC-Fr
Hoechst 33258	Nuclear Stain	Sigma / 94403	1:10,000	IC IC-Fr

Table 2.3 List of antibodies used

Immunocytochemistry (IC), immunocytochemistry on frozen sections (IC-Fr) or flow cytometry (FC)

2.7.2 Tissue and OCT-embedded organoids

Tissue samples were washed twice in HBSS and fixed on ice for 30-45 min in 4% PFA. Tissues were then cryoprotected in 30% sucrose overnight on ice. Organoids were removed from matrigel using cell recovery solution and fixed for 20 min in 4% PFA and washed twice with PBS. Tissue and organoids were then embedded in Optimal Cutting Temperature (OCT), snap frozen, and sectioned at 8 µm. Slides were blocked with 10% donkey serum/0.1% Triton-X for 1 hr. Primary antibodies, diluted in 1% donkey serum/0.1% Triton-X (antibody diluent), were applied overnight at 4°C. Slides were washed 3 times for 5 min with antibody diluent. Secondary antibodies in antibody diluent were applied for 1 hr at RT. Slides were washed three times with antibody diluent for 5 min, nuclei counter-stained with Hoechst 33258 for 5 min, and mounted with Fluoromount G. All images were acquired using either a Zeiss LSM700 laser scanning confocal or a Leica DMLB fluorescent microscope and analysed using ImageJ.

2.8 Quantitative RT-PCR (qPCR)

Total RNA was extracted using either the RNeasy Mini Kit or RNeasy Micro Kit from Qiagen depending on the expected RNA yield of each sample. Total RNA was reverse-transcribed using Superscript II Reverse Transcriptase (Invitrogen). 10µL qPCR reaction mixtures were prepared using the SensiMix SYBR Low-ROX Kit (Bioline) and run on either a Life Technologies QuantStudio 12K Flex or QuantStudio 6 machine in technical duplicate. All genes were normalized to the housekeeping gene Ubiquitin (UBC). Primer sequences used are listed in **Table 2.4**.

2.8.1 Primary human controls for qPCR

Primary human hepatocytes were purchased from Biopredic International, France. RNA was isolated from primary hepatocytes using the GenElute mammalian total RNA isolation kit (Sigma-Aldrich) and qPCR carried out as above in technical duplicate.

Primary human sigmoid colon organoids were derived from biopsy samples and cultured as described previously[124] and RNA isolated using the RNeasy Micro Kit (Qiagen) and qPCR reactions were carried out as above in technical duplicate.

Human embryonic stem cells were cultured as previously described[7] and the line H9 used for all experiments. RNA was isolated using the RNeasy Micro Kit (Qiagen) and qPCR reactions were carried out as above in technical duplicate.

Gene	Forward Primer	Reverse Primer
LGR5	CTCCCAGGTCTGGTGTGTTG	GAGGTCTAGGTAGGAGGTGAAG
PROM1	AGTCGGAAACTGGCAGATAGC	GGTAGTGTTGTACTGGGCCAAT
SOX9	CTCTGGAGACTTCTGAACGAGAG	CCTTGAAGATGGCGTTGGGG
HNF4A	CATGGCCAAGATTGACAACCT	TTCCCATATGTTCCTGCATCAG
ALBUMIN	CCTTTGGCACAATGAAGTGGGTAACC	CAGCAGTCAGCCATTTCACCATAG
TBX3	TGGAGCCCGAAGAAGAGGTG	TTCGCCTTCCCGACTTGGTA
CYP3A4	TGTGCCTGAGAACACCAGAG	GTGGTGGAAATAGTCCCGTG
TTR	ATGGCTTCTCATCGTCTGCT	TGTCATCAGCAGCCTTTCTG
KRT19	ACGACCATCCAGGACCTGC	TCCCACTTGGCCCCTCAGC
KRT7	GATTGCTGGCCTTCGGGGT	TCATCACAGAGATATTCACGGCTC
HNF1B	GCACCCCTATGAAGACCCAG	GGACTGTCTGGTTGAATTGTCG
CDX2	GGCAGCCAAGTGAAAACCAG	TTCCTCTCCTTTGCTCTGCG
SOX17	CGCACGGAATTTGAACAGTA	GGATCAGGGACCTGTCACAC
HOXB2	CCTAGCCTACAGGGTTCTCTC	CACAGAGCGTACTGGTGAAAAA
UBC	ATTTGGGTCGCGGTTCTTG	TGCCTTGACATTCTCGATGGT

Table 2.4 List of primers used for qPCR

2.9 RNA-Sequencing

RNA was isolated from 3 to 7 biological replicates from either organoid cell lines at Passage 5 or from primary extrahepatic tissues as described above. RNA samples were sent for next-generation sequencing at the NIAID Genomics Technologies Branch. Stranded, poly-A purified mRNA truSeq libraries were prepared using the Illumina NeoPrep system. Single-end 75 bp reads were obtained using a NextSeq500. At least 20 million reads/sample evenly distributed across the lanes were obtained. If this number of reads was not reached, the samples were re-sequenced to reach at least 20 million reads. Reads were trimmed to 70 bp and reads less than 40 bp were discarded. Transcript abundance was estimated using *Salmon* with default settings.[125] The *Tximport* package was used to summarize transcript abundances to the gene level and the *DeSeq2* package was used to combine sequencing replicates and for principal component and differential

gene expression analyses in R.[126,127] Only genes classified as protein coding in the Ensembl BioMart GRCh38 database were included in the analyses. A significance cut-off for differentially expressed genes was set as False Discovery Rate (FDR) less than 0.05 and a $\log_2$FoldChange with an absolute value greater than 1 for all analyses. Code used for DeSeq2 analyses in R can be found in **Appendix II.**

Heatmaps of the variance stabilized transformed counts were generated using Morpheus (https://software.broadinstitute.org/morpheus). Hierarchical clustering was performed using one minus pearson correlation with average linkage for all heatmaps. Gene ontology enrichment analyses were performed using GOrilla.[128]

2.10 Statistical Analyses

Statistical analyses were carried out in Prism 7. One-way ANOVAs were performed followed by post-hoc analyses using either Dunnett's, Sidak's, or Tukey's multiple comparisons tests.

3 ESTABLISHMENT AND CHARACTERIZATION OF HUMAN BILIARY ORGANOIDS

3.1 Statement of source

The data and text presented in this chapter are largely based on the following first author manuscript written by the author of this dissertation. Some parts of this chapter have been taken *verbatim* or with only minor changes from this source.

Rimland CA, Tilson S, Morell C, Tomaz R, Lu WY, Adams S, Georgakopoulos N, Otaizo-Carrasquero F, Myers T, Sun HW, Gieseck RL, Sampaziotis F, Tysoe O, Wesley B, Oniscu GC, Hannan NRF, Forbes S, Saeb-Parsy K, Wynn TA, Vallier L. (2018) Regional differences in human biliary tissues and corresponding *in vitro* derived organoids. Manuscript in preparation.

Figures 3.1 and 3.3 contain data from the following publication first published in *Nature Medicine*, on which the author of this dissertation was a co-author and collaborator. However, the data in Figures 3.1 and a portion of the data in Figure 3.3 was generated by Dr. Fotios Sampaziotis and do not represent experiments performed by the dissertation author. These experiments provided proof-of-principle for several of the experiments performed by the author of this dissertation and are presented here to provide important context.

Sampaziotis F, Justin AW, Tysoe OC, Sawiak S, Godfrey EM, Upponi SS, Gieseck RL 3rd, de Brito MC, Berntsen NL, Gómez-Vázquez MJ, Ortmann D, Yiangou L, Ross A, Bargehr J, Bertero A, Zonneveld MCF, Pedersen MT, Pawlowski M, Valestrand L, Madrigal P, Georgakopoulos N, Pirmadjid N, Skeldon GM, Casey J, Shu W, Materek PM, Snijders KE, Brown SE, Rimland CA, Simonic I, Davies SE, Jensen KB, Zilbauer M, Gelson WTH, Alexander GJ, Sinha S, Hannan NRF, Wynn TA, Karlsen TH, Melum E, Markaki AE, Saeb-Parsy K, Vallier L. (2017). Reconstruction of the mouse extrahepatic biliary tree using primary human extrahepatic cholangiocyte organoids. *Nature Medicine. 23(8): 954-963.*

Three of the liver tissue samples used for derivation of intrahepatic bile duct organoids were obtained in collaboration with Dr. Wei-Yu Lu and Dr. Stuart Forbes at the University of Edinburgh. Additionally, three of the intrahepatic organoid cell lines grown in extrahepatic organoid conditions were isolated and maintained in culture by Dr. Wei-Yu Lu. However, all characterization (qPCR, IF analyses, etc.) on intrahepatic organoids was performed by the author of this dissertation.

3.2 Introduction

The biliary tree is a series of interconnected ductular tissues responsible for the drainage, storage, and concentration of bile produced by the liver and pancreatic juices from the pancreas, which are released into the intestine to aid in digestion.[47] The biliary tree can be broadly divided into two main compartments, that which is within the liver, intrahepatic, and that which is outside of the liver, extrahepatic.[47,48] The intrahepatic compartment is further divided based on duct size, while the extrahepatic compartment is divided by anatomical region, including the common hepatic duct, gallbladder (GBD), cystic duct, common bile duct (CBD).[48] Additionally, the main pancreatic duct (PancD) is often considered as part of the extrahepatic biliary tree.[48] Embryologically, the extrahepatic and intrahepatic biliary compartments arise from different precursor cells.[74] Intrahepatic cholangiocytes arise from bipotent hepatoblasts, which also give rise to hepatocytes. Extrahepatic cholangiocytes share an embryologic origin with the ventral pancreas.[74]

Research on cholangiocyte diversity has historically focused exclusively on the intrahepatic compartment.[56] Analysis of intrahepatic bile ducts in rodents has demonstrated the existence of two populations of cholangiocytes: small and large

cholangiocytes.[55,57] Small cholangiocytes line the smallest intrahepatic ductules and are cuboidal in shape. Large cholangiocytes line larger intrahepatic ducts and are columnar in shape. Further, it has been shown that large and small cholangiocytes within the intrahepatic compartment display differences in transcriptional profiles, proliferative capacity, and biological function.[55,57] Despite the identification of heterogeneity within the liver, research has not yet been extended to characterization of the extrahepatic compartment. Little is known on what diversity, if any, may exist between individual anatomic regions of the extrahepatic biliary system. This basic understanding is critical as many human cholangiopathies such as biliary atresia, primary sclerosing cholangitis, and non-anastomotic biliary strictures preferentially target extrahepatic bile ducts (EHBDs) early in disease.[84,85,129]

Additionally, it has recently been demonstrated that *in vitro* 3D-culture of human biliary organoids is possible from both intrahepatic and extrahepatic tissues.[97,98,116,122] However, it remains unclear whether the extrahepatic cultures, like other epithelial organoid systems such as intestinal organoids, contain true stem/progenitor cell populations.[93] Huch et al (2015) isolated human intrahepatic bile ducts (IHBDs) and cultured them in conditions promoting canonical WNT signalling. These IHBD organoids were capable of long-term expansion and bi-potential differentiation capacity towards either hepatocyte or biliary cell fates.[97] This work strongly suggested the presence of an intrahepatic stem/progenitor cell population. Using a similar WNT based culture system, Lugli et al (2016) established both mouse and human extrahepatic GBD organoids.[116] However, the differentiation capacity was only assessed for the murine organoids and showed incomplete differentiation towards a hepatocyte fate with no *in vivo* differentiation potential demonstrated.[116] Further, our group recently demonstrated that human biliary organoids from CBD and GBD, grown in the presence of a canonical WNT pathway inhibitor, were capable of long-term expansion, yet lacked characteristic markers of adSCs such as LGR5 and PROM1.[122] The differentiation capacity of these organoids beyond a biliary fate was not assessed in our prior work.[122] Taken together, these reports suggest that both intra- and extrahepatic human cholangiocytes are capable of forming 3D-organoid cultures, which under certain culture conditions display adult tissue stem/progenitor phenotypes. However, it is still unclear whether EHBD organoids are capable of bi-potential differentiation like IHBD organoids. Additionally, how EHBD organoids from different anatomical regions of the biliary tree compare to each other, to their tissue of origin, and to IHBD organoids is also unknown.

In this chapter, we describe the derivation and characterization of human biliary organoids from three major regions of the human extrahepatic biliary tree: CBD, GBD, and PancD. We demonstrate that these EHBD organoids, when grown in media conditions promoting canonical WNT signalling, express markers of adSCs such as LGR5 and PROM1. We additionally establish IHBD organoids and show they are divergent in culture condition requirements from EHBD organoids.

In chapter 4, we use these organoids and primary tissue samples in order to perform RNA-sequencing to assess what differences exist between (1) the three extrahepatic tissue regions, (2) extrahepatic tissues compared to their corresponding organoid cultures, (3) intrahepatic organoids compared to extrahepatic organoids, and (4) intrahepatic tissues compared to extrahepatic tissues.

Lastly, in chapter 5 we assess the differentiation capacity of both EHBD and IHBD organoids towards a hepatocyte fate.

3.3 Results

3.3.1 Mechanical dissociation allows for enrichment of human extrahepatic biliary epithelial cells

We previously reported a method for the isolation of biliary epithelial cells using mechanical dissociation.[122] This method results in a higher viability than enzymatic dissociation methods and enriches for cells co-expressing the biliary markers KRT7 and KRT19 **(Figure 3.1)**. Given our previous results, we employed this method again to isolate biliary epithelial cells from human CBD, GBD, and PancD.

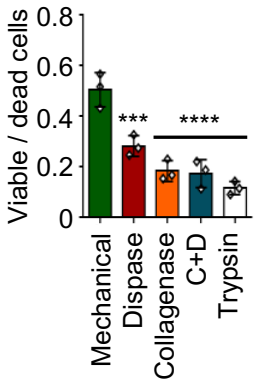


Figure 3.1 Mechanical dissociation of biliary epithelial cells improves viability.

Ratio of viable to dead primary biliary epithelial cells after being subjected to different methods of dissociation. One-way ANOVA with Dunnett correction comparing mechanical dissociation to others, ***= p<0.001, ****=p<0.0001. Error bars are standard deviation and each dissociation method was attempted three independent times (n=3). Data and figure generated by Dr. Fotios Sampaziotis. Reproduced with permission from Sampaziotis et al (2017)[122] first published in *Nature Medicine*. C= collagenase. D= dispase.

Our previous work demonstrated mechanical dissociation from CBD donors (n=3) yielded 94.6% ± 2.4% (mean ± SD) of cells co-expressing KRT7/KRT19.[122] Since this method had only been assessed in extrahepatic ductal tissue samples, we re-validated this method of isolation on epithelial scrapings from GBD donors (n=2), which showed a similar percentage of cells co-expressing KRT7/KRT19 using mechanical dissociation (94.7% ± 4.31%, mean ± SD) (**Figure 3.2 and 3.3**). Established organoid cultures also demonstrated a high level of purity with an average of 95.04 ± 2.09 % (n=5) of cultured cells co-expressing KRT19/KRT7 (**Figure 3.2 and 3.3**).

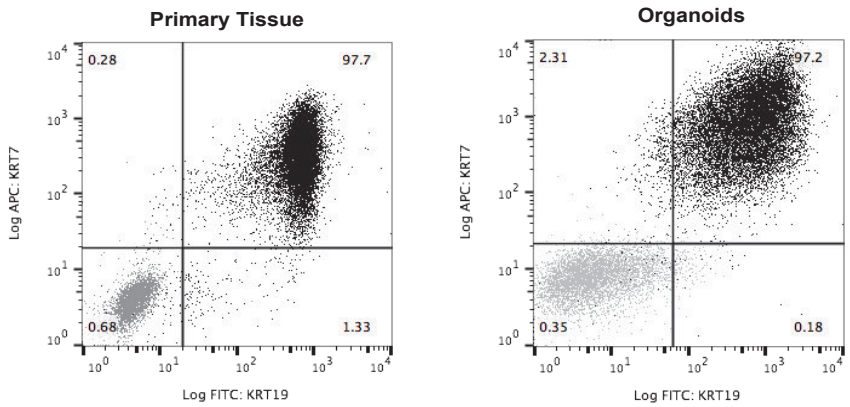


Figure 3.2 Flow cytometry analysis for KRT7 and KRT19

Representative flow cytometry analysis of primary epithelial cells mechanically dissociated from human gallbladder tissue (n=2) or *in vitro* cultured extrahepatic biliary organoids (n=5) showing percentage of cells co-expressing KRT7 and KRT19 (black). Tissue and organoids represent independent experiments and gates were set for each experiment separately. Negative control: secondary antibody only stained cells (grey). Single stained/fluorescence minus one and secondary only stained controls were used to set gates.

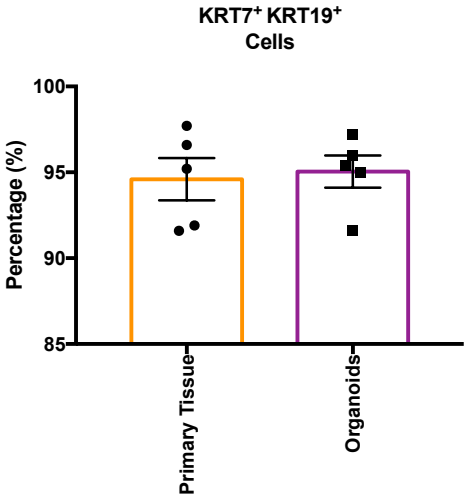


Figure 3.3 Percentages of cells co-expressing KRT7 and KRT19

Percentage of cells co-expressing KRT7 and KRT19 determined by flow cytometry analysis of either epithelial cells mechanically dissociated from primary human gallbladder tissue (n=2), epithelial cells mechanically dissociated from primary human common bile duct tissue (n=3), or *in vitro* cultured extrahepatic biliary organoids (n=5 cell lines). Flow cytometry data for the three common bile duct primary tissue samples was generated by Dr. Fotios Sampaziotis.[122] All other data points were generated by the author of this dissertation.

3.3.2 Extrahepatic biliary epithelial cells can be cultured as 3D organoids in conditions promoting canonical WNT signalling

In the course of our previous experiments, we noted that extrahepatic cholangiocytes cultured as organoids in the presence of RSPO-1, DKK-1 (a canonical WNT pathway inhibitor), and EGF were capable of long-term expansion but did not express characteristic adSC markers such as LGR5 and PROM1.[122] In parallel, we also observed that growing these extrahepatic cholangiocyte organoids in the presence of CHIR 99021, a small molecule inhibitor of GSK3β, instead of DKK-1, resulted in lower levels of phosphorylated ß-Catenin and thus a higher level of WNT signalling activity.[122] Given these observations, we decided to test a new culture media which promotes canonical WNT signalling. We hypothesized this media would allow for the isolation of extrahepatic biliary cells with a progenitor phenotype.

We based this new culture media on protocols previously used by our group to maintain human intestinal organoids derived from induced pluripotent stem cells.[130] This media consisted of R-Spondin 1, CHIR 99021, Noggin, PGE_2, and EGF. We found this media allowed for establishment of EHBD organoid cultures, which expressed markers of adSCs such as LGR5 and PROM1. However, all cultures eventually deteriorated, and proliferation ceased 4-12 weeks after isolation, such that cultures could only be maintained at a 1:1-1:2 splitting ratio.

Around this time, a report was published by Huch et al. (2015) on human IHBD organoids and demonstrated that for long-term proliferation, addition of a TGF-β inhibitor (A 83-01) and cAMP activator (Forskolin, FSK) was necessary.[97] Given this finding, we attempted to rescue our deteriorating EHBD organoids by addition of 5 μM A 83-01 and 10 μM FSK (**Figure 3.4**). Within 48 hours, addition of A 83-01 and FSK to late passage organoids rescued the cultures. Splitting ratios returned to 1:3 or greater. The growth impact of A 83-01 and FSK was also noticeable even when added at early passages before proliferation deteriorated (**Figure 3.4**). Addition of A 83-01 and FSK did not significantly change the expression profile of the organoids and they continued to express adSC markers and biliary markers equal to our original media conditions (**Figure 3.5**).

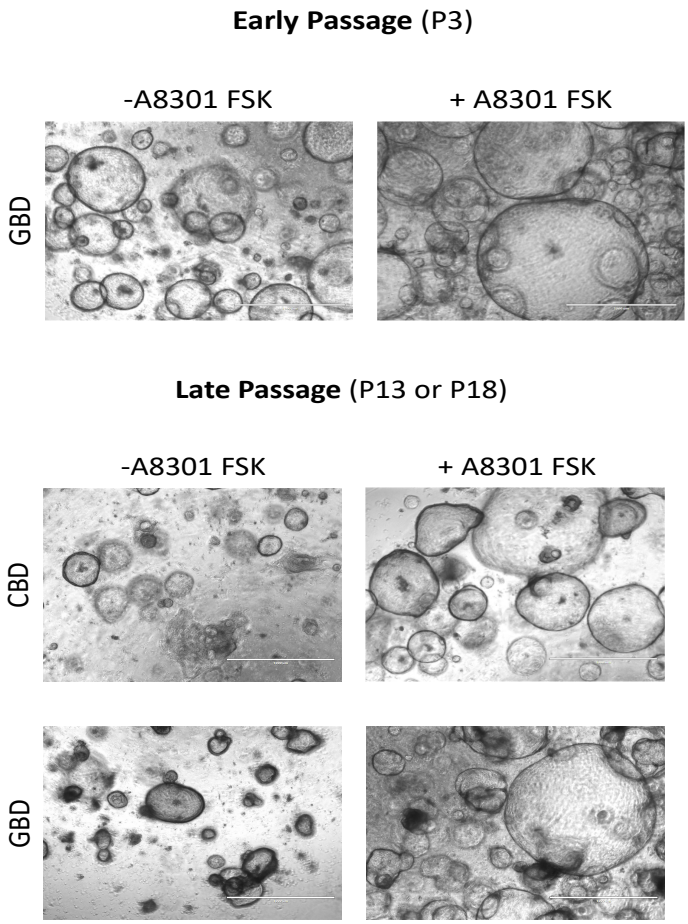


Figure 3.4 Images of organoids cultured with or without A 83-01 and FSK.

Scale bar = 1000 μm

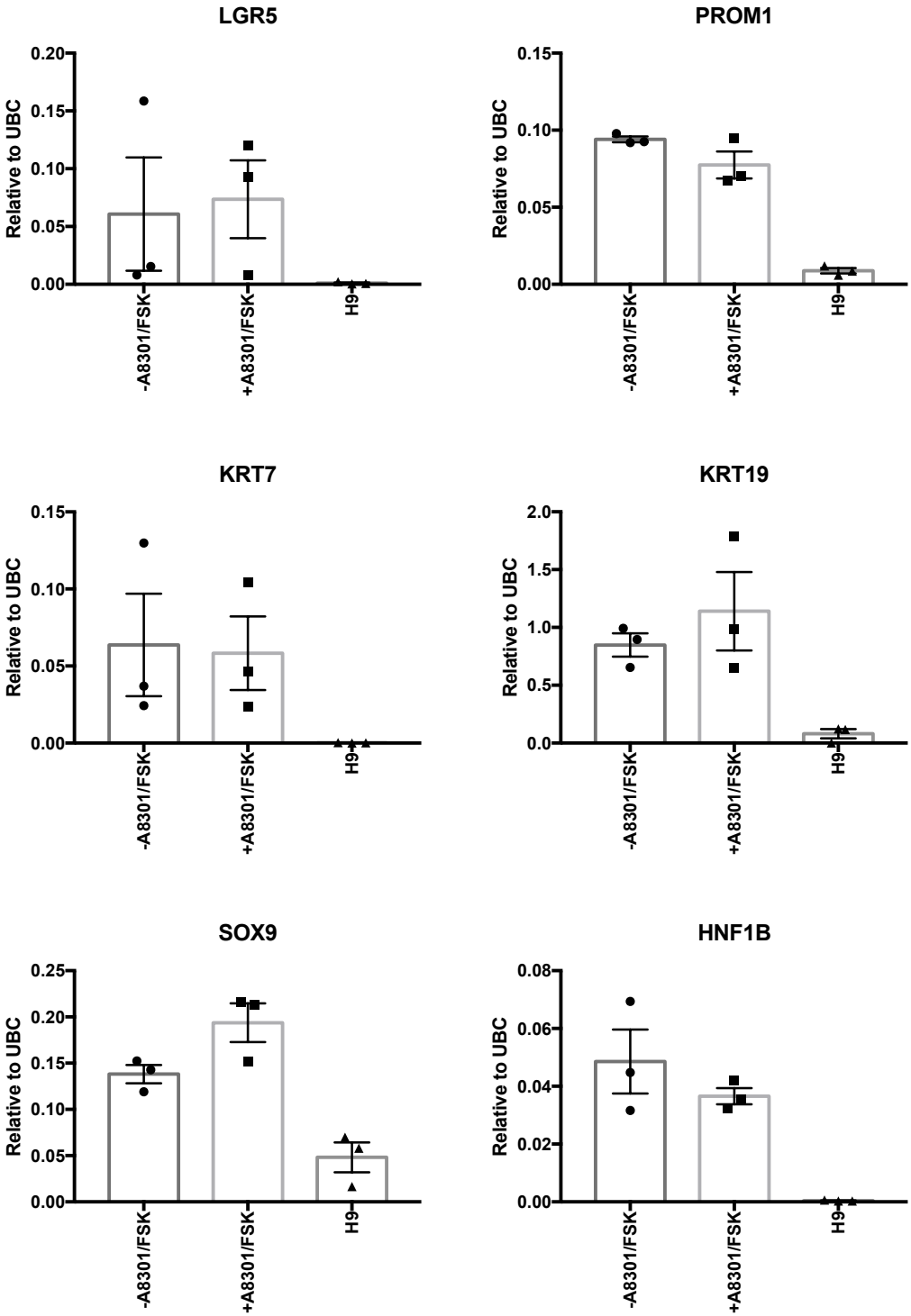


Figure 3.5 qPCR analysis of extrahepatic bile duct organoids cultured with or without A 83-01 / FSK.

Extrahepatic organoids were cultured with or without A8301 and FSK (n=3). Unpaired t-tests comparing +A8301/FSK vs. -A8301/FSK were not significant for any genes ($p > 0.05$). Gene expression is normalized to the housekeeping gene UBC and data is plotted as mean and SEM. H9 embryonic stem cells (n=3) were used as negative controls.

All future tissue samples were derived in these new optimized conditions and in total we derived 7 CBD lines, 9 GBD lines, and 3 PancD lines in these conditions from donors ranging in age from 20-68 years old with 100% of the tissue samples yielding organoid cultures. The organoids displayed similar morphology and growth dynamics regardless of their tissue of origin and could not be readily distinguished from one another (**Figure 3.6**).

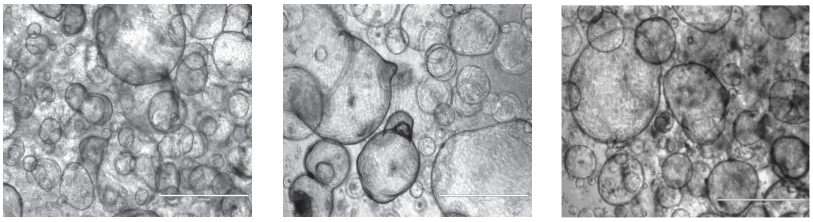


Figure 3.6 Images of organoids isolated from three regions of the extrahepatic biliary tree

Representative images of organoids isolated from CBD, GBD, and PancD extrahepatic tissues.

Scale Bars: 1000 μm

3.3.3 Extrahepatic organoids express markers of adult stem cells while maintaining biliary markers and histological similarity to primary tissues

With our culture conditions stabilized, we sought next to characterize the organoids derived from each region of the extrahepatic biliary tree. qPCR analyses demonstrated that EHBD organoids expressed adSC markers (LGR5 and PROM1), ductal markers (SOX9, KRT7, HNF1B), and early hepatocyte markers (TBX3, HNF4A), but not markers of mature hepatocytes (ALB) (**Figure 3.7**) and that most of these markers remained consistently expressed over time in culture (Passage 0 vs. Passage 10+, **Figure 3.8**). Levels of ductal markers in EHBD organoids were comparable to or greater than primary tissue. With the exception of the mature biliary marker, HNF1B, which was reduced in the EHBD organoids. The stem/progenitor cell markers LGR5 and PROM1, and TBX3, were significantly enriched in EHBD organoids compared to primary tissue. Expression of all markers by qPCR did not differ significantly between CBD, GBD, and PancD primary tissues. Furthermore, CBD, GBD, and PancD organoids did not differ significantly for any of the markers examined.

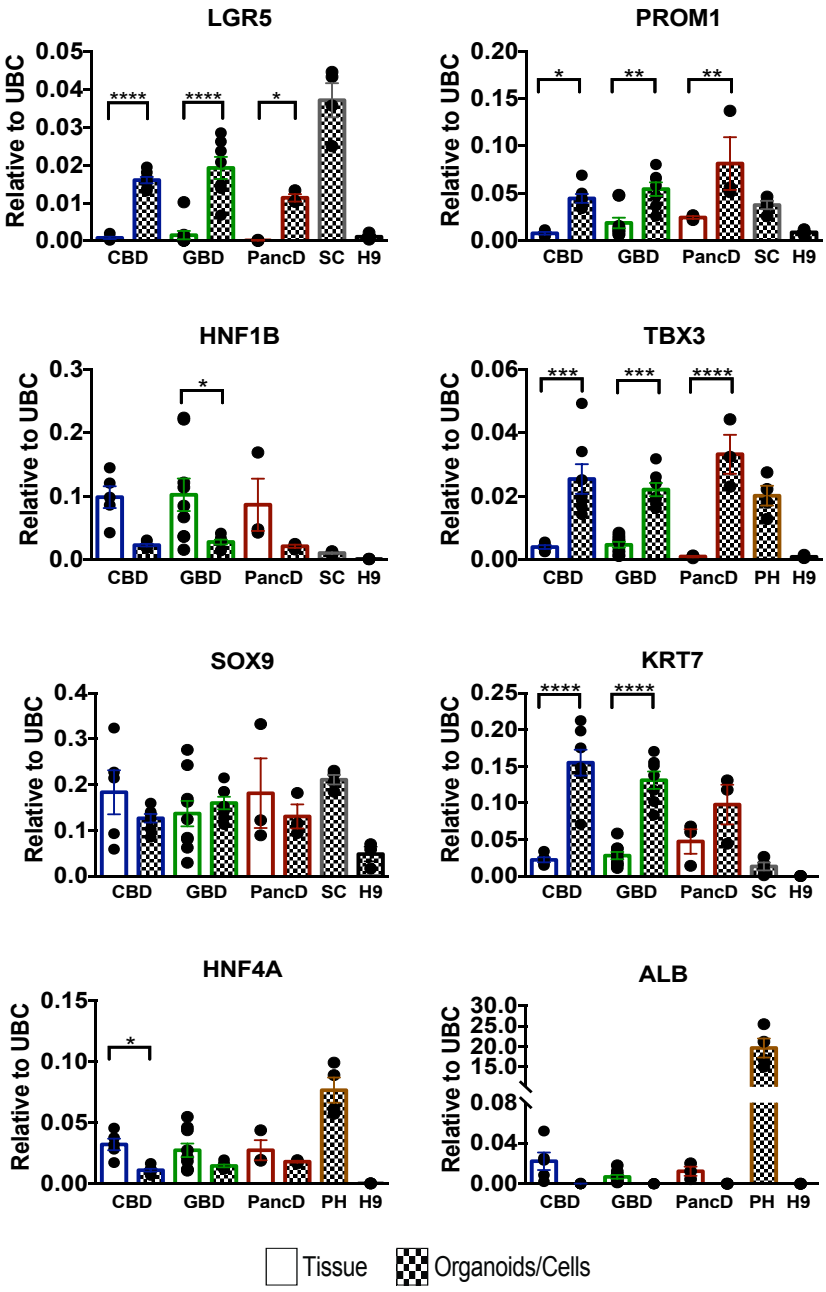


Figure 3.7 qPCR characterization of extrahepatic bile duct organoids and primary tissues demonstrates expression of adult stem cell and biliary markers

qPCR analyses showing the expression of biliary and adSC markers in EHBD organoids and primary tissue samples from common bile duct (CBD), gallbladder (GBD), and pancreatic duct (PancD). Organoids derived from human sigmoid colon biopsies (SC, n=4), primary hepatocytes (PH, n=4, 2 donors), and H9 human embryonic stem cells (H9, n=3) were used as controls. Gene expression is normalized to the housekeeping gene UBC and data is plotted as mean and SEM. CBD Tissue (n=5), CBD organoids at P5 (n=7), GBD tissue (n=9), GBD organoids at P5 (n=7), PancD tissue (n=3), PancD organoids at P5 (n=3). *=$p \leq 0.05$, **=$p \leq 0.01$, ***=$p \leq 0.001$, ****$p \leq 0.0001$. If not otherwise indicated, comparisons between groups were not significantly different ($p>0.05$).

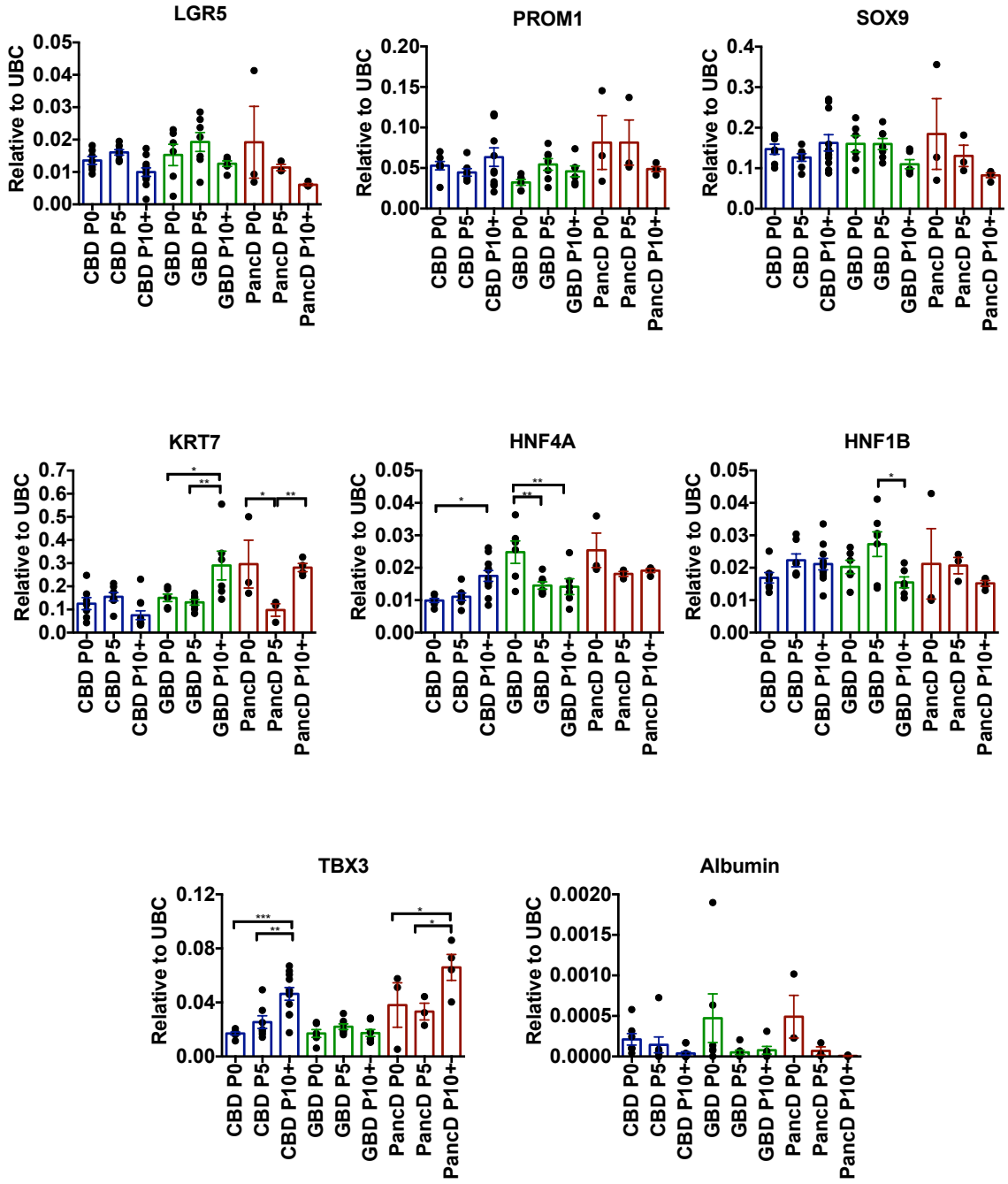


Figure 3.8 qPCR characterization of extrahepatic bile duct organoids over time in culture demonstrates maintanence of biliary and adult stem cell markers.

qPCR analyses showing the expression of biliary and adSC markers in EHBD organoids (n=3-11, n=2-7 donor lines per passage) derived from common bile duct (CBD), gallbladder (GBD), and pancreatic duct (PancD) tissues over time in culture (P0, P5, P10 or higher). Gene expression is normalized to the housekeeping gene UBC and data is plotted as mean and SEM. *=$p \leq 0.05$, **=$p \leq 0.01$, ***=$p \leq 0.001$, ****$p \leq 0.0001$. If not otherwise indicated, comparisons between groups were not significantly different ($p>0.05$).

Immunofluorescence (IF) analyses demonstrated a substantial fraction of cells were actively proliferating as indicated by KI67 positivity (**Figure 3.9**). Interestingly, primary tissues had no detectable levels of KI67, suggesting that proliferation is extremely rare in the biliary epithelium. Finally, the organoids bore a striking similarity to their tissues of origin, with both organoids and tissues uniformly expressing EPCAM, SOX9, KRT19, HNF4A and KRT7.

Taken together, these results demonstrate biliary organoids can be derived from any region of the extrahepatic biliary tree and that the addition of WNT induces the expression of stem cell markers in these cell lines. These extrahepatic bile duct (EHBD) organoids can be distinguished from our previously described organoids cultured without canonical WNT, which did not express stem cell markers and instead more closely resembled fully differentiated cholangiocytes.

3.3.4 Intrahepatic bile duct organoids cannot be maintained in extrahepatic culture conditions and display divergent characteristics from EHBD organoids.

We next sought to establish organoids from intrahepatic bile ducts (IHBD) using the same culture conditions used to derive EHBD organoids. For that, liver tissue obtained from cadaveric donor was subjected to enzymatic dissociation and the resulting cells grown in the culture conditions described above for extrahepatic organoids. The resulting organoids (IHBD_ex) could be established and grew well for the first 2-3 weeks in culture. Interestingly, the IHBD_ex organoids were very different from the EHBD organoids (**Figure 3.10**) and displayed a heterogeneous morphology with some organoids having a cryptic phenotype with budding structures, while others maintained a cystic structure similar to EHBD organoids.

After 5-7 passages, these cryptic structures began to predominate, proliferation ceased and cultures could not be maintained. Given this result, we also isolated IHBD organoids in conditions previously published by Huch and colleagues (2015).[97] The resulting organoids (IHBD_Huch) grown in these culture conditions maintained their proliferative ability for at least 10 passages without evidence of deterioration (**Figure 3.11**) while displaying a homogenous cystic morphology.

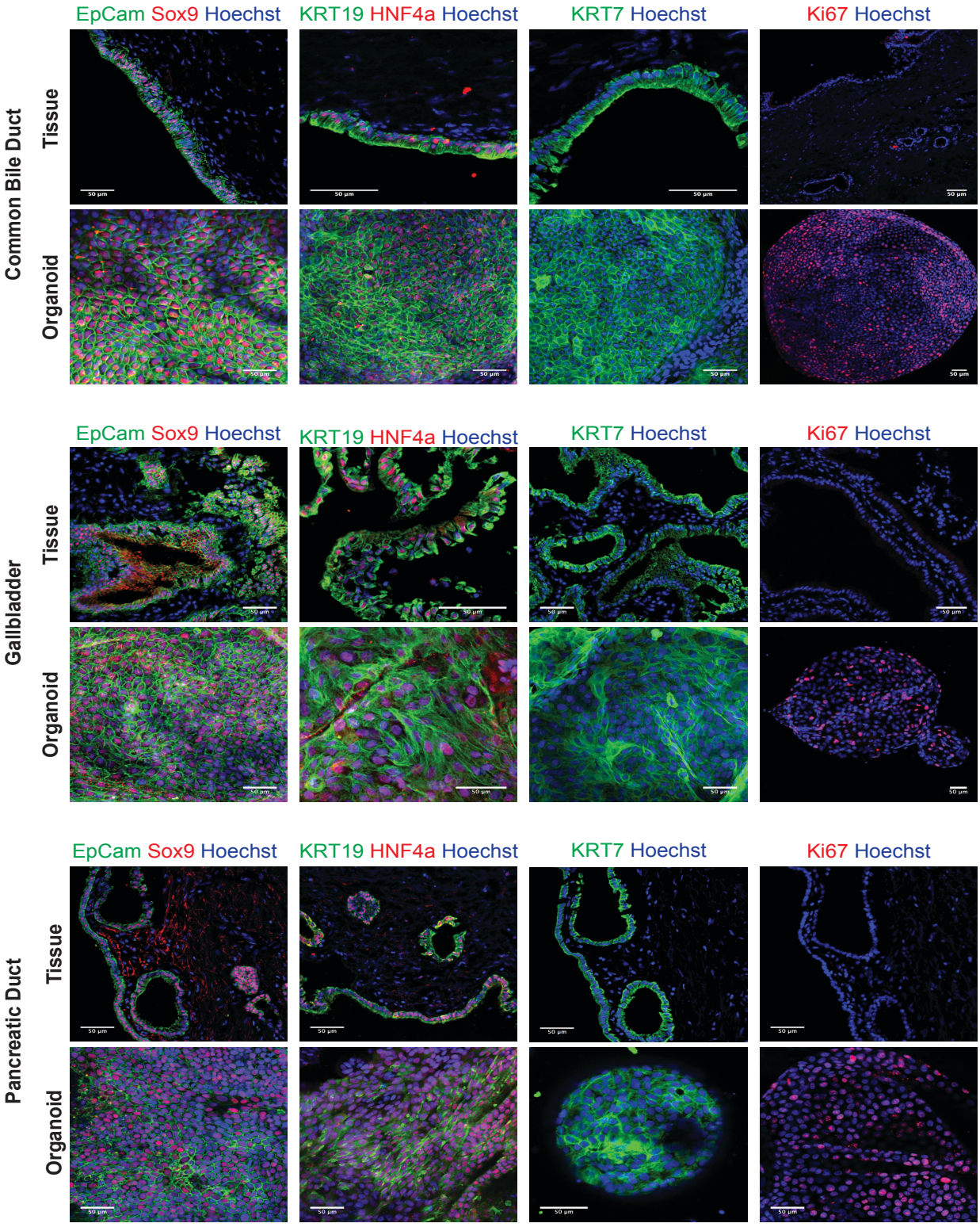


Figure 3.9. Extrahepatic primary tissues and corresponding organoids express biliary markers and show histologic similarities.

Representative immunofluorescence images of extrahepatic bile duct organoids and primary tissues for selected markers. Scale bars: 50 µm.

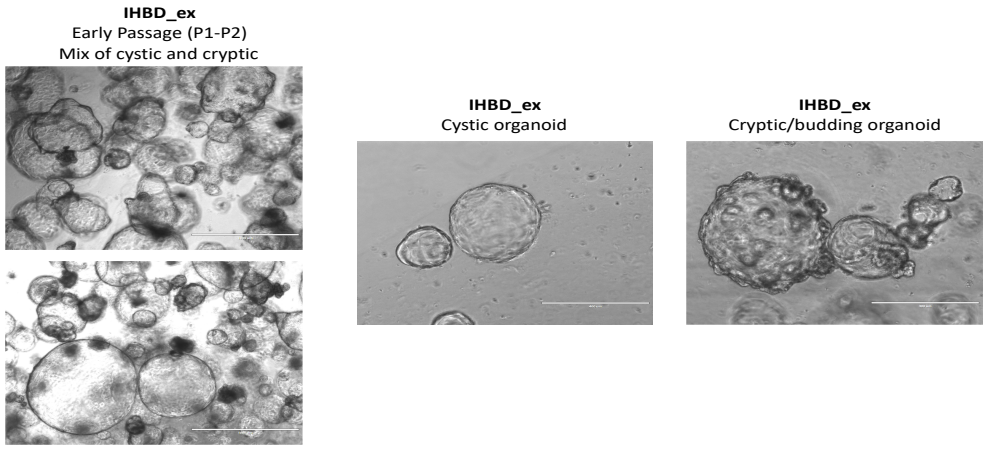


Figure 3.10 Intrahepatic bile duct organoids cultured in extrahepatic media conditions display morphological heterogeneity.

Images of intrahepatic bile duct organoids cultured in extrahepatic culture conditions (IHBD_ex) highlighting the morphological heterogeneity of the organoids with both cystic and budding/cryptic organoids present in the cultures. Scale bar = 1000 or 400 μm

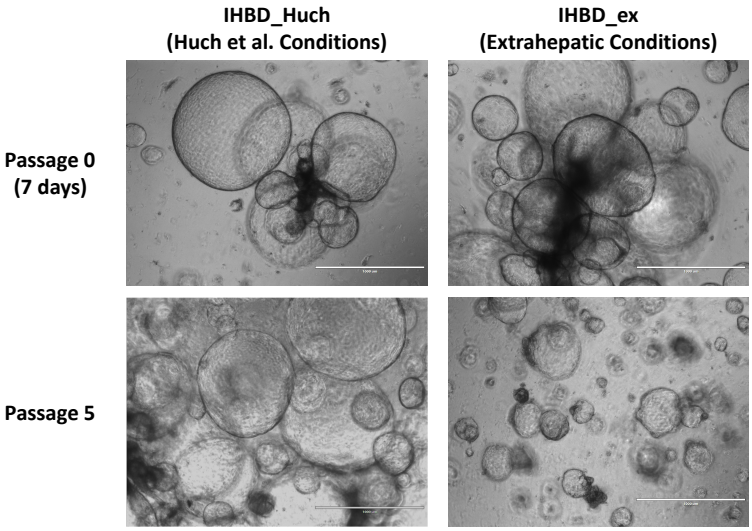


Figure 3.11 Intrahepatic bile duct organoids cultured in extrahepatic media conditions cannot be maintained long-term in culture.

Images of intrahepatic bile duct organoids cultured in either extrahepatic culture conditions (IHBD_ex) or Huch et al (2015) conditions (IHBD_Huch) at Passage 0 or Passage 5. Scale bar = 1000 μm.

qPCR analyses between EHBD, IHBD_ex and IHBD_Huch organoids revealed only minor differences in the expression of biliary markers and stem cells markers (**Figure 3.12**). IF confirmed that, similarly to EHBD organoids, IHBD organoids expressed EPCAM, SOX9, KRT19, and KRT7 (**Figure 3.13**). Of note, IHBD organoids expressed HNF4A like their EHBD counterparts, but this transcription factor was absent from primary IHBD tissue. KI67 expression was present in a substantial number of cells in IHBD_Huch organoids. While this marker for proliferative cells was variable in the IHBD_ex organoids and the budding/cryptic organoids demonstrated fewer KI67 positive cells than cystic organoids in the same culture (**Figure 3.14**). KI67 was absent in the primary tissue (**Figure 3.13**). The morphological variation between the budding/cryptic and the cystic IHBD_ex organoids was also readily apparent (**Figure 3.14**).

Further investigation into these morphologic and proliferative differences revealed that the presence of CHIR 99021 in the culture media was most likely responsible for the divergence between conditions (**Figure 3.15**). IHBD_Huch organoids were transferred into various media combinations including extrahepatic organoid conditions with or without CHIR 99021 or alternatively into the conditions described by Huch and colleagues (2015) with or without CHIR 99021. Only organoids in the conditions with CHIR 99021 demonstrated the budding morphology. Additionally, conditions with CHIR 99021 present showed decreased organoid size, fewer organoids, and could not be maintained for more than 4 passages. Whereas conditions without CHIR 99021 continued to proliferate.

We also performed qPCR for adult stem cell, biliary, and hepatocyte specific genes on the IHBD organoids in these culture conditions with or without CHIR 99021 hoping to gain a better understanding of why such morphological and proliferative differences were observed (**Figure 3.16**). However, for the genes we examined, very few significant changes in expression were observed. Suggesting that for these genes at least, CHIR 99021 does not significantly alter the expression profile in the organoids.

Taken together, these data demonstrate that cells from the extra- and intrahepatic biliary tree diverge in their growth factor requirements with continuous WNT signalling being detrimental for the proliferation and maintenance of IHBD organoids but not for EHBD organoids. However, this difference seems to have little consequence on the expression of biliary markers.

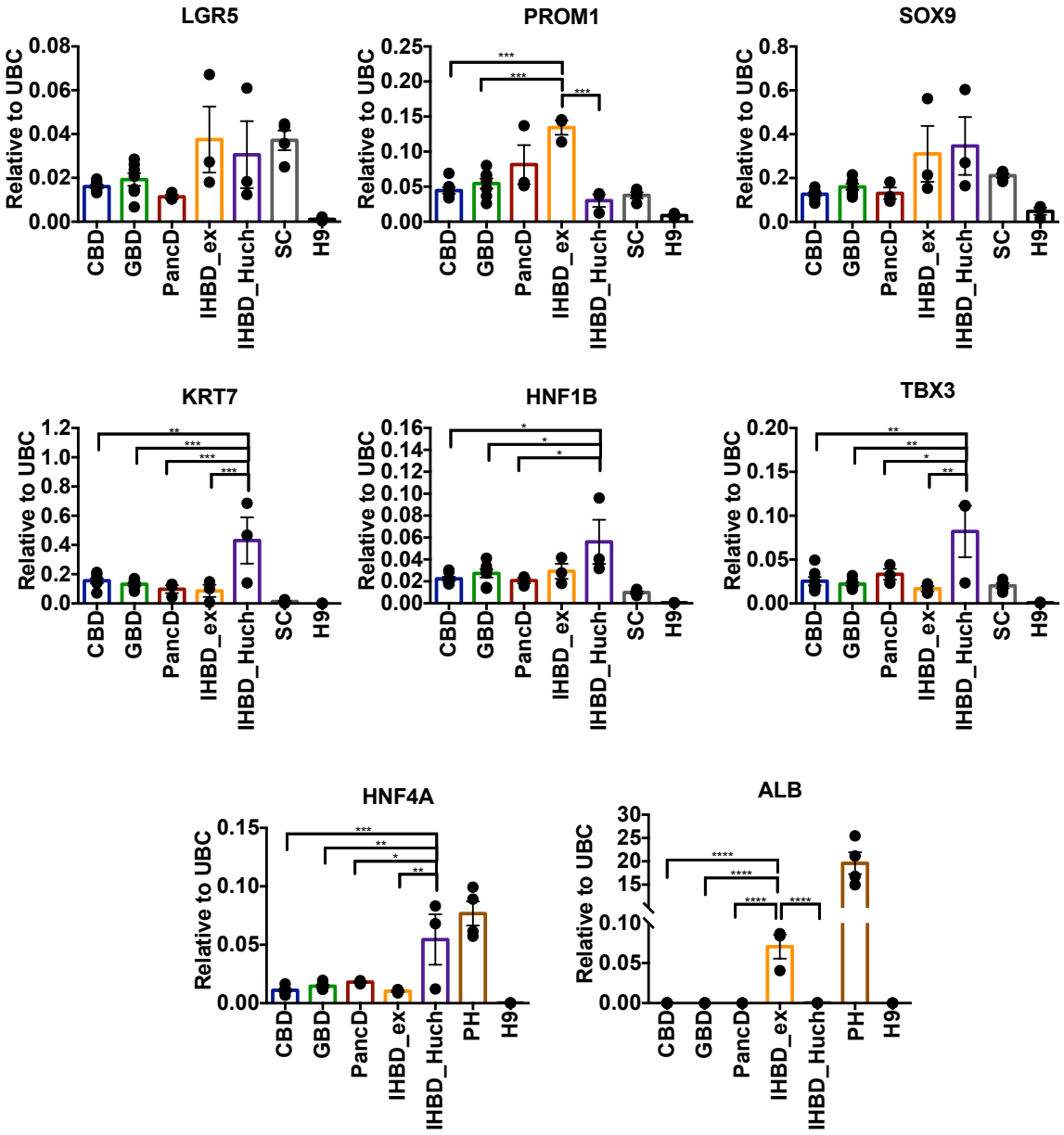


Figure 3.12 qPCR analysis of intrahepatic and extrahepatic bile duct organoids demonstrates subtle differences for selected markers.

qPCR analysis of IHBD_ex (n=3) and IHBD_Huch (n=3) organoids in comparison to extrahepatic organoids from common bile duct (CBD) (n=7), gallbladder (GBD) (n=7), and pancreatic duct (PancD) (n=3) at passage 5. Primary hepatocytes (PH, n=4), sigmoid colon organoids (SC, n=4), and H9 embryonic stem cells (H9, n=3) were used as controls. Gene expression is normalized to the housekeeping gene UBC and data is plotted as mean and SEM. IHBD_ex organoids were derived in collaboration with Dr. Wei-Yu Lu and Dr. Stuart Forbes. *=p ≤ 0.05, **=p ≤ 0.01, ***=p ≤ 0.001, ****p ≤ 0.0001. If not otherwise indicated, comparisons between groups were not significantly different (p>0.05)

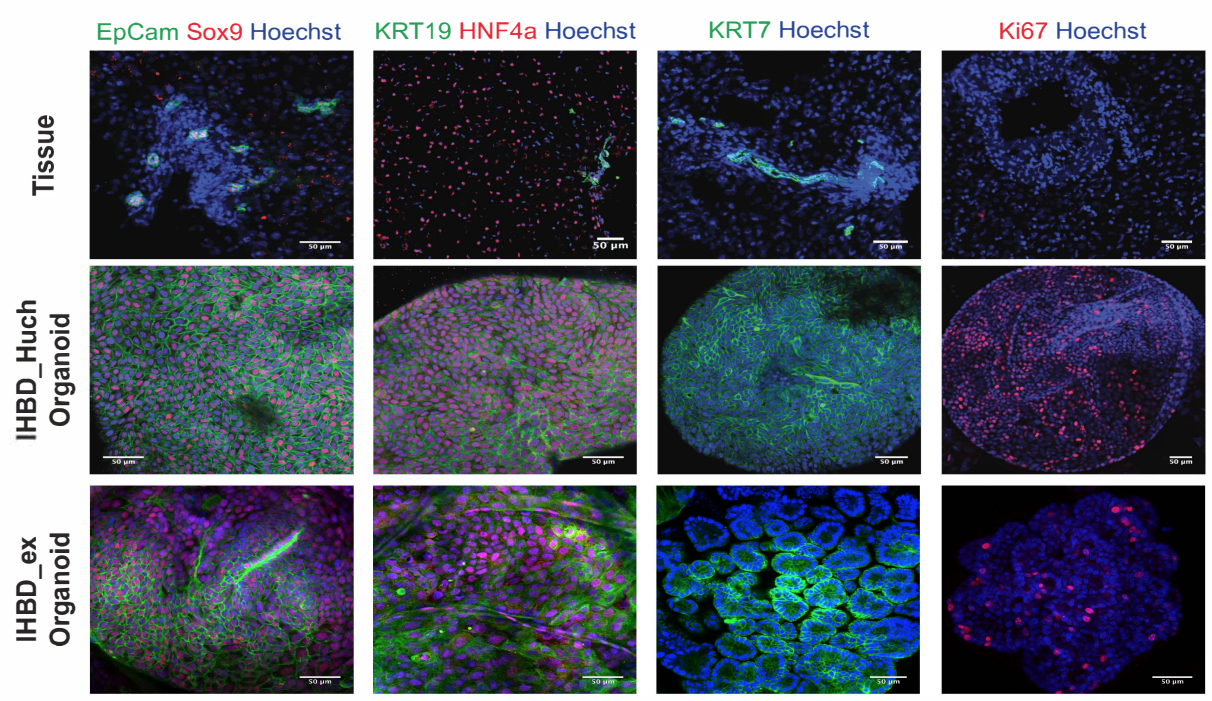


Figure 3.13 Intrahepatic bile duct primary tissues and corresponding organoids express biliary markers and show histologic similarities.

Immunofluorescence analysis for select markers of IHBD_Huch, IHBD_ex, and primary liver tissue. Scale Bars: 50 µm

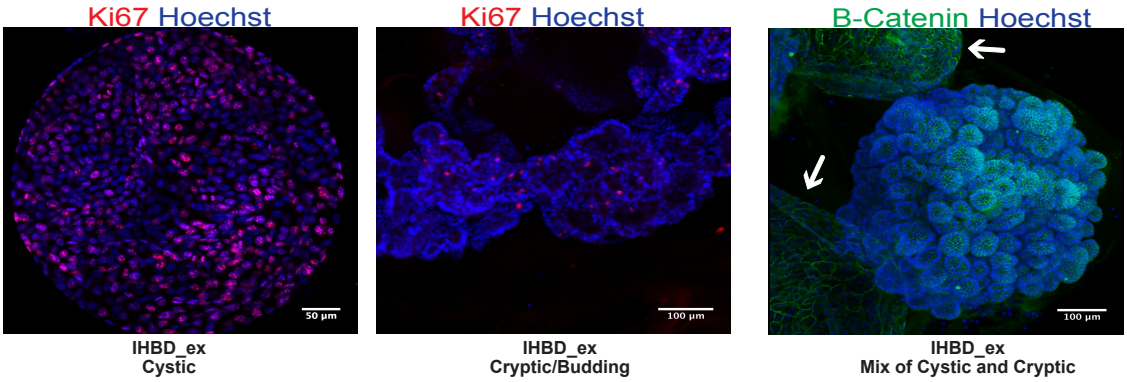


Figure 3.14 Immunofluorescence analysis of IHBD_ex organoids highlights the morphological heterogeneity of the organoids.

Immunofluorescence analysis for KI67 and ß -Catenin on IHBD_ex organoids. White arrows indicate cystic organoids. Scale bar=50-100 µm

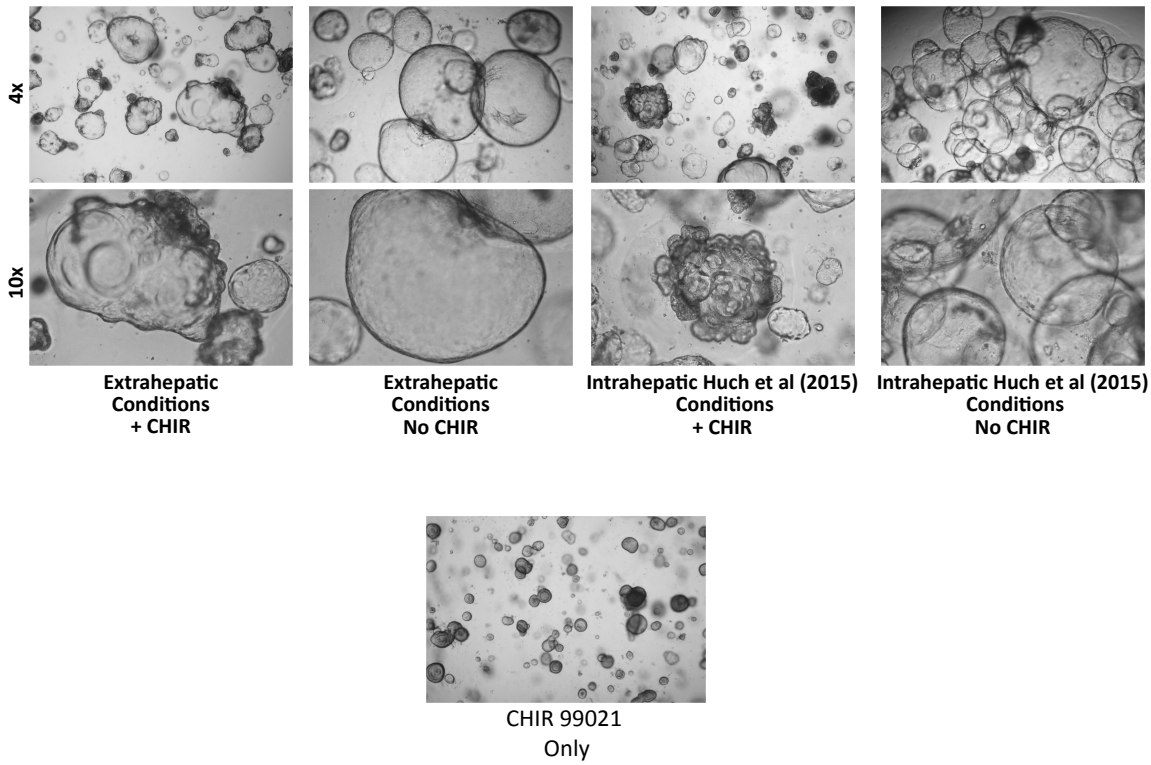


Figure 3.15 IHBD_Huch organoids treated with CHIR 99021

Intrahepatic bile duct organoids isolated and cultured originally in conditions published by Huch et al (2015) were transferred into different medias with or without 3 uM CHIR 99021 added. The five media combinations included: EHBD organoid media with CHIR 99021, EHBD organoid media without CHIR 99021, IHBD_Huch organoid media with CHIR 99021, IHBD_Huch media without CHIR 99021, or media with only CHIR 99021. Within 72 hr, morphological changes could be observed, and budding/cryptic structures developed in organoids treated with CHIR 99021 similar to what had previously been observed in IHBD organoids isolated and cultured in extrahepatic organoid conditions which contained CHIR 99021.

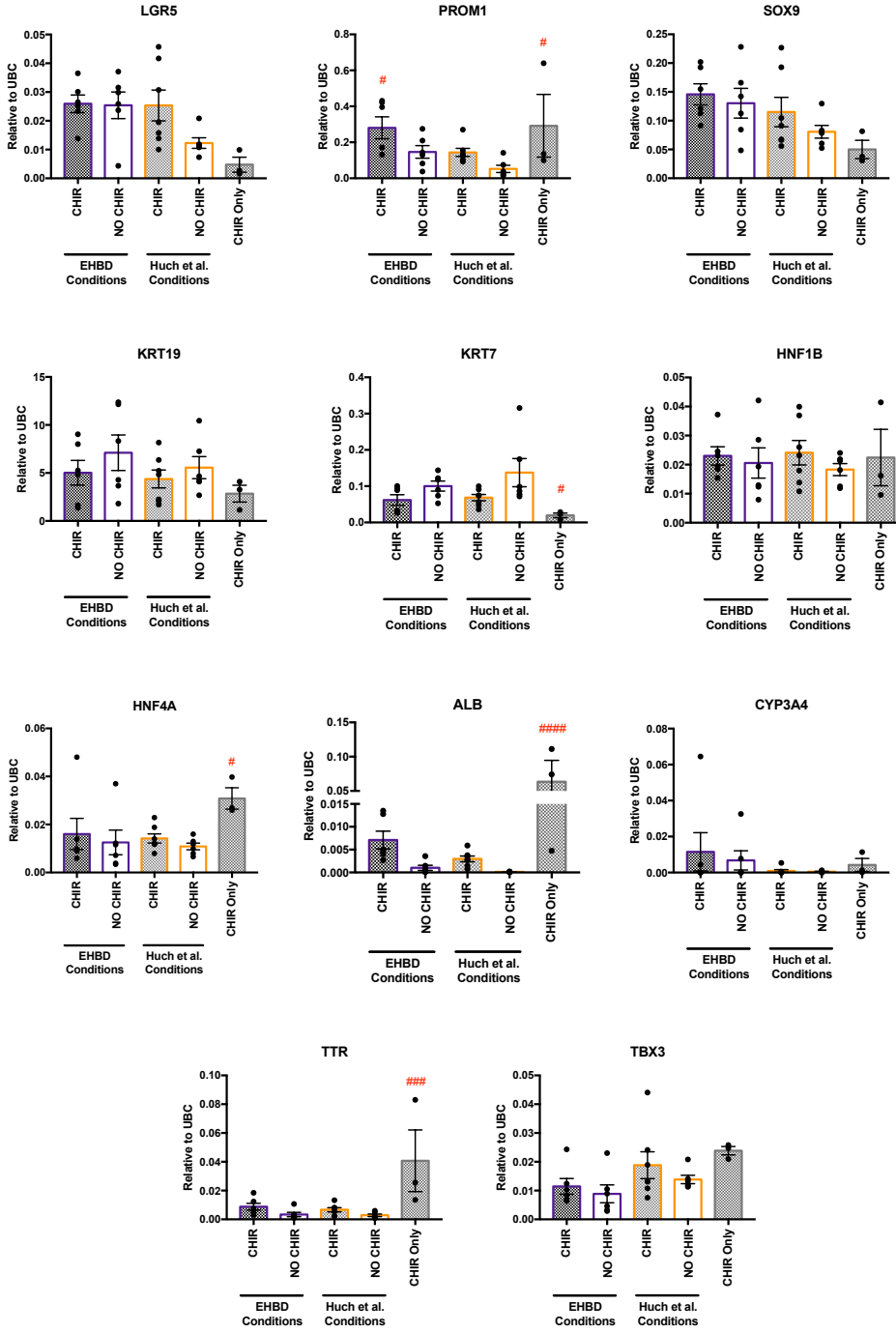


Figure 3.16 qPCR analysis of IHBD_Huch organoids cultured in Huch et al (2015) conditions or EHBD organoid conditions with or without CHIR 99021.

qPCR analysis of IHBD_Huch organoids (n=3 donor lines) transferred into various media conditions to assess the effect of CHIR 99021 on the expression profile of organoids. Conditions included EHBD organoid media with CHIR 99021 (n=6), EHBD organoid media without CHIR 99021 (n=6), Huch et al conditions with CHIR 99021 (n=7), Huch et al conditions without CHIR (n=6) and Basal Media with only CHIR 99021 (n=3). # = $p \leq 0.05$, ## = $p \leq 0.01$, ### = $p \leq 0.001$, #### = $p \leq 0.0001$ for the particular media condition compared to Huch et al Conditions NO CHIR. If not otherwise indicated, comparisons between groups were not significantly different ($p>0.05$)

3.4 Discussion

In this chapter, the derivation and characterization of biliary organoids from four regions of the human biliary tree including: GBD, CBD, PancD, and IHBDs was described. To do this, culture conditions that promote canonical WNT signalling were developed allowing for the isolation and propagation of cells with a stem/progenitor cell phenotype. In the first portion of this chapter, the enrichment of extrahepatic biliary epithelial cells from extrahepatic tissues through the use of mechanical dissociation was described. Using these mechanically dissociated cells, it was shown that EHBD organoids from CBD, GBD, and PancD can be derived with a 100% efficiency rate and from a large number of donors and donors of various ages. These extrahepatic organoids show significantly higher expression of adult stem cell markers LGR5 and PROM1 when compared to their primary tissue of origin. The EHBD organoids also display long-term proliferation capacity suggesting these organoid cultures may represent a stem/progenitor population. Lastly, in the second half of the chapter, the derivation of organoids from intrahepatic bile duct tissue was described. It was discovered that IHBD organoids cannot be maintained long-term in the same conditions used for EHBD organoids. Further, when cultured in extrahepatic conditions, IHBD derived organoids are morphologically very different from EHBD organoids and this may be due to CHIR 99021. Together, these results suggest key differences exist between IHBD and EHBD organoids.

We feel the work presented in this chapter represents a significant progression for the field. To our knowledge, it is the first attempt to directly compare biliary organoids derived from these different anatomical regions of the biliary tree. However, there are limitations to the work, which must be acknowledged.

The first limitation is that we were unable to culture both IHBD and EHBD derived organoids long-term in the same media conditions. When we first began this work, the report by Huch et al (2015) on the derivation of IHBD organoids had not been published.[97] Therefore, when beginning our attempts to derive EHBD organoids in WNT-based conditions for the first time, we used protocols already established in our lab. These conditions allowed us to establish EHBD organoids and our early characterizations suggested enrichment for a stem/progenitor phenotype. We did not anticipate that our EHBD organoids would cease proliferating, as many of the cell lines we derived grew well for more than 1-2 months in culture. However, as we continued to derive additional EHBD organoids we noticed the organoids could not be maintained and began to consider

that our culture conditions were sub-optimal. Around this time, the report by Huch et al (2015) was published and reported the need for A 83-01 and FSK. As we had already put a considerable amount of work in to the early characterization of our EHBD organoids, we decided to assess whether adding A 83-01 and FSK into our existing conditions would rescue the cultures and it was able to do so. Adding A 83-01 and FSK to the EHBD organoids did not change their expression profile significantly and the organoids continued to be enriched for stem/progenitor markers. It is interesting that for long-term proliferation, the organoids required inhibition of TGF-ß signalling. TGF-ß has been implicated in cholangiocyte quiescence/senescence and has been shown to act as a key negative regulator of cholangiocyte proliferation.[131,132]

Given these results, we moved forward with this culture system and performed the characterization experiments described in this chapter. After this, we then attempted to derive IHBD organoids. We quickly realized that IHBD organoids in extrahepatic conditions (IHBD_ex) are unable to proliferate long-term. To circumvent this limitation, we also isolated IHBD cells in the culture conditions reported by Huch et al (2015) (IHBD_Huch) and characterized these alongside of the IHBD_ex to account for differences in proliferation. EHBD organoids were not derived in Huch et al (2015) conditions as doing so would have required a large number of new donor tissue samples, which were unavailable to us at the time. Although it would have been ideal to have IHBD and EHBD organoids cultured in the same media conditions this was not possible. It is important to note, that despite this being a limitation, it did allow us to discover that IHBD and EHBD organoids are intrinsically different and their requirement for different media conditions suggests these organoid systems are not equivalent to one another. Further, we found preliminary results suggesting that response to CHIR 99021/WNT signalling may be responsible for the divergence between IHBD and EHBD organoid culture conditions. IHBD_Huch organoids adopted morphological heterogeneity when CHIR 99021 was present in the media. Interestingly, this result is similar to what has been previously observed with human gastric organoids and may suggest a similar mechanism acts in IHBD organoids.[20] However, these results are preliminary and should be explored further in future studies.

An additional limitation is that our mechanical dissociation method for EHBD tissue does not isolate biliary epithelial cells with 100% purity. An alternative to these isolation methods would have been to use cell sorting and to isolate cells from each tissue which express markers such as KRT19, KRT7, or EPCAM for instance. However, in order to

use such methods, cells must be dissociated to single cells and enzymatic dissociation greatly reduced the viability of extrahepatic biliary epithelial cells making this challenging. Additionally, the number of cells obtained from common bile duct, and especially pancreatic duct tissues, was often low. It was already challenging to obtain enough cells from these tissues for primary tissue RNA, let alone enough viable cells for sorting. We also found that the contamination by non-biliary cell types is likely low with only 5-10% of the cells isolated by mechanical dissociation not expressing KRT19/KRT7. Overall, we felt that mechanical dissociation, although having these obvious limitations, was the best approach to enrich for EHBD epithelial cells. Further, it has been well-established in the literature that organoid culture systems such as these are selective for epithelial cells, making it unlikely that our cultures contain non-epithelial cells and this was supported with our flow cytometry and IF analyses. This has also been previously demonstrated for enzymatically dissociated, manually picked IHBD ducts and IHBD organoid cultures are only derived from ductal cells and not from other liver cell types.[97]

Lastly, despite our characterizations, it is still unclear whether our EHBD organoids represent a stem/progenitor population. Although informative, qPCR and IF characterizations are limited in the number of markers which can be reasonably assessed. To address some of these limitations, the following two chapters attempt to further characterize the nature of the organoids described in this chapter. In chapter 4 we expand our characterization to include the entire transcriptome through the use of RNA-Sequencing (RNA-Seq) profiling. Finally, in chapter 5 we explore the differentiation capacity of EHBD and IHBD organoids in order to further assess the stem/progenitor characteristics possessed by these biliary organoids.

4 TRANSCRIPTOMIC PROFILING OF HUMAN BILIARY TISSUES AND *IN VITRO* ORGANOIDS

4.1 Statement of Source

The results presented here are largely based on experiments presented in the following first author manuscript written by the author of this dissertation. Therefore, some parts have been taken *verbatim* or with only minor changes from this source.

Rimland CA, Tilson S, Morell C, Tomaz R, Lu WY, Adams S, Georgakopoulos N, Otaizo-Carrasquero F, Myers T, Sun HW, Gieseck RL, Sampaziotis F, Tysoe O, Wesley B, Oniscu GC, Hannan NRF, Forbes S, Saeb-Parsy K, Wynn TA, Vallier L. (2018) Regional differences in human biliary tissues and corresponding *in vitro* derived organoids. Manuscript in preparation.

The RNA-Sequencing experiments presented in this chapter were performed in collaboration with a number of individuals.

The RNA samples for IHBD_ex organoids and IHBD tissues were generated in collaboration with Drs. Wei-Yu Lu and Stuart Forbes. The NIAID Genomics Technologies Branch at the NIH, particularly Dr. Timothy Meyers and Mr. Francisco Otaizo-Carrasquero, performed the library preparations and sequencing. The mapping and transcriptome-based alignment using *Salmon* was performed by Dr. Rute Tomaz.

Differential gene expression analyses in R using *DeSeq2,* and all other analyses, including figure preparation, were performed by the author of this dissertation. Dr. Rute Tomaz and Dr. John Ferdinand provided support, example code, and guidance while performing the differential gene expression analyses.

4.2 Introduction

RNA-Sequencing has gained in popularity over the last 10 years as a method for profiling the entire RNA transcriptome of a biological sample.[133,134] The transcriptome represents all of the RNA transcripts present in a given biological sample at a point in time and includes messenger RNA (mRNA), non-coding RNAs, and small RNAs. Profiling of the entire transcriptome allows for the comparison of individual transcript expression levels across biological groups and can provide insight into differential expression patterns. Originally, transcriptome analyses were only accomplished using low-throughput methods such as PCR or northern blots. These techniques only allowed for the profiling of a small number of transcripts at a time. With the advent of microarray technology, the transcriptome could be assayed for the first time. Microarray technology has limitations such as a need to know the transcript sequence beforehand and difficulties assessing the expression of transcripts with highly similar sequences. RNA-Sequencing eliminates these limitations.

The general principle behind RNA-Sequencing is similar to methods employed for high-throughput DNA sequencing but cDNA libraries are prepared from reverse transcribed RNA samples (**Figure 4.1**). First, ribosomal RNA is depleted using either poly-adenosine selection or ribosomal depletion. The RNA is then reverse transcribed into cDNA, fragmented, and chemical adapters ligated to each sample for identification. Samples are placed into lanes on a flow cell and a high throughput sequencer is used for sequencing. Most high throughput sequencers use sequencing by synthesis chemistry, which works by incorporating fluorescently labelled nucleotides complementary to the cDNA fragment. These fluorescent nucleotides are recorded by the sequencer and output as a FASTQ file. From the sequencing FASTQ file, millions of reads corresponding to the sequences present in the fragmented cDNA samples are obtained.

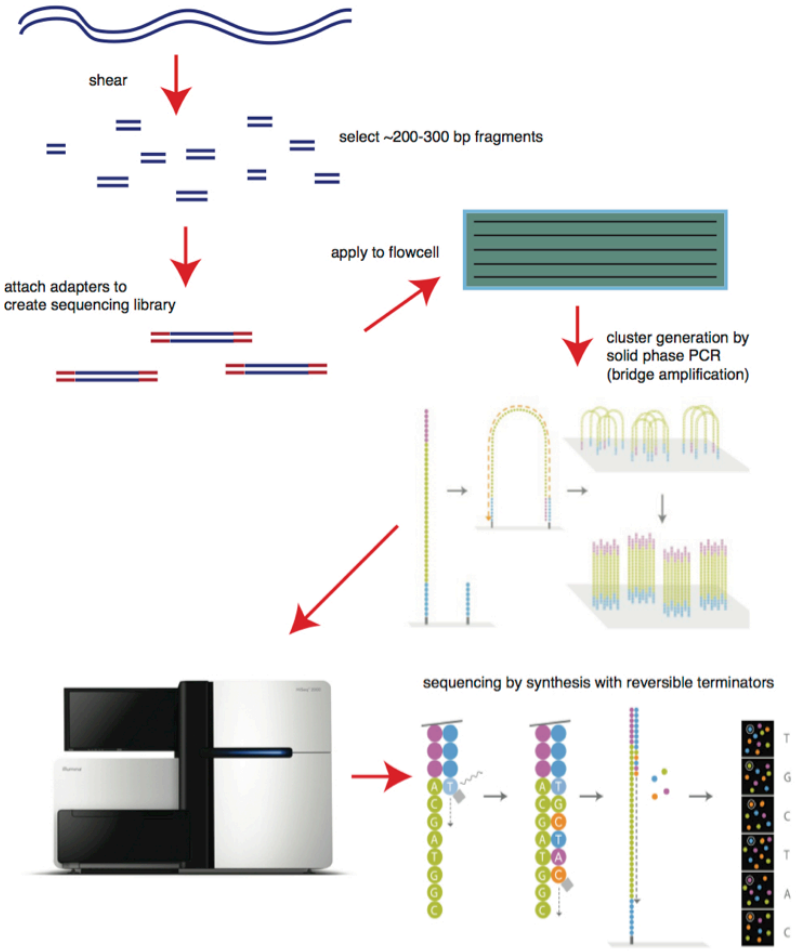


Figure 4.1 Diagram depicting the workflow for performing RNA-Seq.

Reproduced from http://web.bioinformatics.cicbiogune.es/

These reads are then aligned to either the genome or transcriptome and the mapped reads are counted to estimate expression levels. This count data is then normalized to account for differences in sequencing depth between samples and used to perform differential gene expression analyses using statistical models such as *DeSeq2* or *EdgeR* that takes advantage of the negative binomial distribution of RNA-Seq data.[127,135] Ultimately, a list of differentially expressed genes between biological groups can be obtained. These gene lists can be used for further analyses, for example, biological process/pathway enrichment analyses such as gene ontology, gene set enrichment analysis, or pathway analysis.[128,136,137] These analyses provide insight into potential biologically meaningful differences between groups and can be used for biomarker discovery.

In this chapter, we take advantage of RNA-Sequencing technology to perform differential gene expression analysis on EHBD primary tissues, EHBD organoids, IHBD organoids,

and IHBD tissue samples. To our knowledge, only human gallbladder tissue has been previously sequenced. As part of the Human Protein Atlas effort, human gallbladder donor tissue samples were sequenced and the transcriptome compared to 27 other tissues.[138,139] Only five genes were found to be enriched in gallbladder tissue, when defined as having five-fold higher expression in gallbladder tissue than all other tissues. This included the genes FGF19, CHST4, MOGAT1, UGT2B28, and AC083862.1. A further 135 genes were identified which were either "group-enriched" (five-fold higher in gallbladder tissue and at least one other tissue type versus the others) or "enhanced" (genes that were five-fold higher in gallbladder tissue in comparison to the average expression level across all other 27 tissues). This work represents an important effort in characterizing the transcriptome of one region of the human extrahepatic biliary tree. However, these experiments were performed on whole GBD tissue, not epithelial enriched tissue samples. Further, other regions of the extrahepatic biliary tree have not been sequenced and it remains unclear how the transcriptome of these tissues compare to each other.

4.3 Results

4.3.1 RNA isolation from human biliary tissue and in vitro organoids

In order to perform our RNA-Sequencing experiments, we first isolated RNA from mechanically dissociated, epithelial enriched, human CBD, GBD, and PancD tissue samples. We also isolated RNA from enzymatically dissociated and manually picked IHBD ducts. RNA was also isolated from CBD, GBD, PancD, IHBD_ex, and IHBD_Huch organoids. RNA quantity and quality were assessed on an agilent bioanalyzer (**Table 4.1**). It was found that EHBD tissue samples had some level of degradation with RNA integrity numbers (RIN) ranging from 4.1-8.9. RIN is a measure of the degradation present in RNA samples and has values from 1 (completely degraded) to 10 (intact). In examining the literature, it seems this may be an intrinsic challenge with these tissues, especially when using mechanical dissociation.[140] As we wanted to match the primary tissue and cell line donors whenever the amount of starting material would allow, it was decided not to re-attempt RNA isolation for the EHBD tissue samples. Additionally, obtaining a large number of donor samples for a second time was not feasible. Therefore, we continued with the sample set and performed sequencing experiments in collaboration with the NIH NIAID Genomics Technologies Branch.

4.3.2 Batch effect and outlier assessment of RNA-sequencing data

Due to sample availability, we were required to sequence our samples in two separate batches (**Table 4.1**). Additionally, thirteen samples in the first batch did not have enough read depth after the first round of sequencing. Those thirteen samples required deeper sequencing and were re-sequenced with our second batch of samples. This created three potential batches within the data. Batch 1A consisted of samples which had library preparation performed in 2016 (libbatch = 1) and were sequenced in 2016 (seqyear = A). Batch 1B consisted of samples which had library preparation in 2016 (libbatch = 1) but then required deeper sequencing in 2017 (seqyear = B). Batch 2B consisted of samples which became available after the initial sequencing, with both library preparation and sequencing performed in 2017 (libbatch = 2, seqyear = B). Additionally, when our second batch of samples was being prepared, new libraries were prepared of four samples from the initial libbatch = 1 set to serve as internal batch controls. These four replicates were then re-sequenced alongside batch 2B samples. Before moving forward with down-stream analyses of our RNA-Seq data, the presence of batch effects in the data needed to be assessed. As seen in **Figures 4.2 and 4.3,** there were no observable batch effects. For the analyses described below, data was combined from technical replicates across sequencing runs and data from 7 CBD tissue, 5 GBD tissue, 3 PancD tissue, 6 CBD organoid, 3 GBD organoid, 3 PancD organoid, 4 IHBD_ex organoid, 3 IHBD_Huch organoid, and 3 IHBD tissue samples was ultimately analysed.

In the early, exploratory analyses of the data, it was noted that one of the IHBD_ex organoid samples appeared to be an outlier from the other three samples (**Figure 4.4**). This particular sample was from an IHBD_ex organoid line at passage 3 and not passage 5 like the others. We included it in the analyses to increase our number of samples. However, when the sample was collected, the cells were still moderately proliferative, unlike the other lines at passage 5 whose growth had slowed. Given this known variability in the samples we excluded this IHBD_ex sample from further analyses.

Sample	Internal ID	RIN	Donor	Batch
CBD Tissue 1	CBD4	5.7	1	1A
CBD Tissue 2	CBD5/PD4	5.6	2	1A / 1B
CBD Tissue 3	CBD8	4.7	4	1A / 1B
CBD Tissue 4	PD1	6.5	19	1A
CBD Tissue 5	PD2	5.6	20	1A
CBD Tissue 6	PD3	5.7	21	1A
CBD Tissue 7	PD5	5.5	22	1A/1B/2B
GBD Tissue 1	GBD4	4.2	7	1A
GBD Tissue 2	GBD6	4.1	8	1A
GBD Tissue 3	GBD8	4.4	1	1A/1B/2B
GBD Tissue 4	GBD10	4.2	5	1A/1B
GBD Tissue 5	GBD11	4.7	3	1A/1B
PancD Tissue 1	PancD1	8.9	15	2B
PancD Tissue 2	PancD2	8.6	16	2B
PancD Tissue 3	PancD5	6.7	18	2B
IHBD Tissue 1	H01	6.3	10	1A/1B
IHBD Tissue 2	H02	7.3	11	1A/1B
IHBD Tissue 3	H03	7.1	12	1A/1B
CBD Organoids 1	CBD4	8.4	1	1A
CBD Organoids 2	CBD5/PD4	9.6	2	1A
CBD Organoids 3	CBD6	9.5	3	1A / 1B
CBD Organoids 4	PD2	9.8	20	1A
CBD Organoids 5	PD3	9.7	21	1A
CBD Organoids 6	PD5	9.5	22	1A/1B/2B
GBD Organoids 1	GBD8	9.8	1	1A
GBD Organoids 2	GBD10	8.7	5	1A
GBD Organoids 3	GBD11	9.5	3	1A/1B/2B
PancD Organoids 1	PancD1	8.2	15	2B
PancD Organoids 2	PancD2	10	16	2B
PancD Organoids 3	PancD4	9.3	17	2B
IHBD_ex Organoids 1	IHBD1_c	10	10	2B
IHBD_ex Organoids 2	IHBD2_c	9.3	11	2B
IHBD_ex Organoids 3	IHBD3_c	10	12	2B
IHBD_ex Organoids 4	IHBD5_c	9.7	14	2B
IHBD_Huch Organoids 1	IHBD1_m	9.8	10	2B
IHBD_Huch Organoids 2	IHBD4_m	9.6	13	2B
IHBD_Huch Organoids 3	IHBD5_m	10	14	2B

Table 4.1 Details on RNA integrity number, donor, and batch for samples sequenced.

RIN = RNA Integrity number. Batch 1A = samples library prepped and sequenced in 2016. Batch 1B= samples library prepped in 2016 that required deeper sequencing in 2017. Batch 2B = samples library prepped and sequenced in 2017.

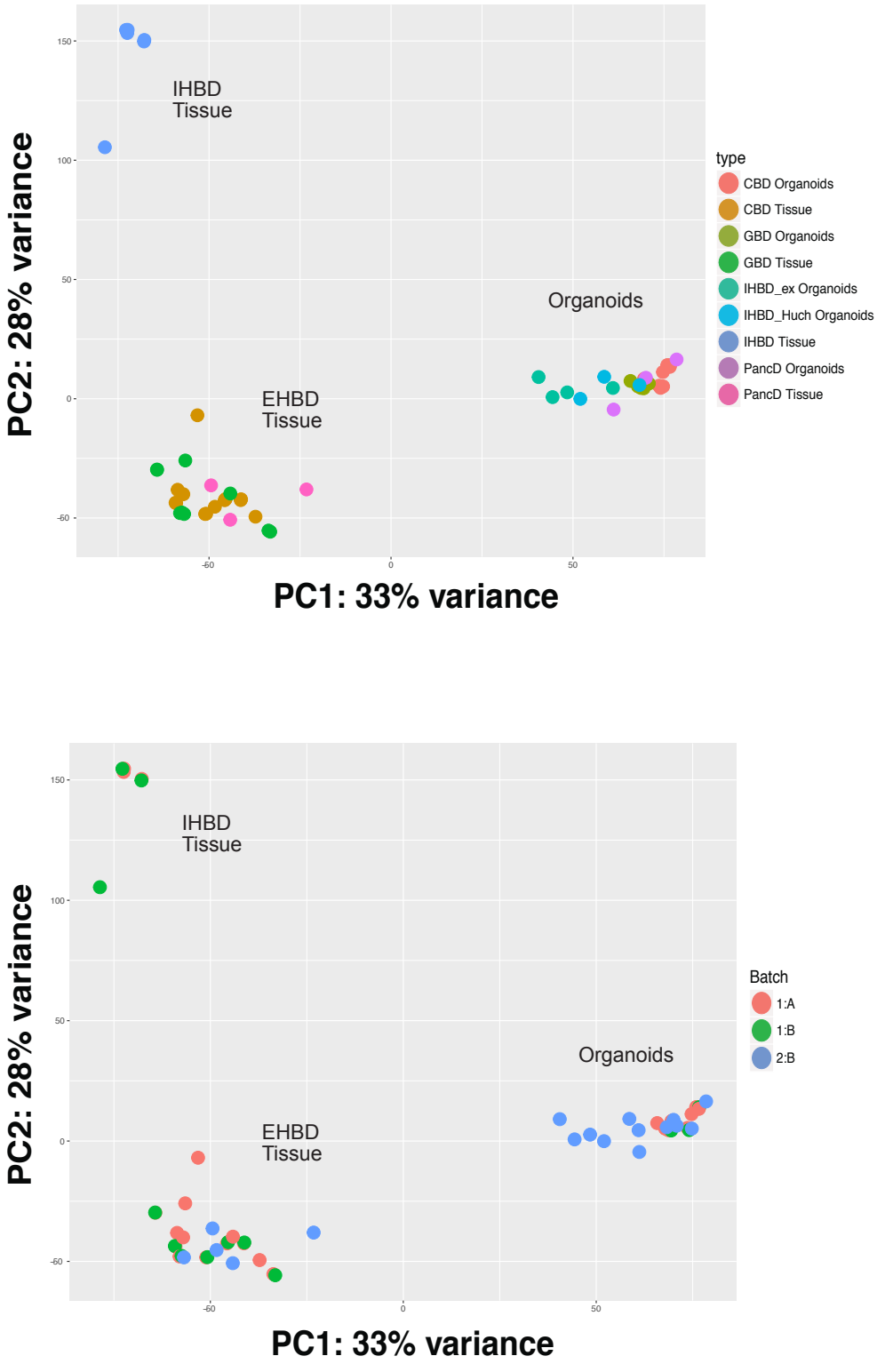


Figure 4.2 Principal component analysis of all samples to assess for batch effects.

Principal component analysis of the *DeSeq2* variance stabilized counts for all genes with greater than zero counts across all samples (19,376) with samples labelled by either type or by batch.

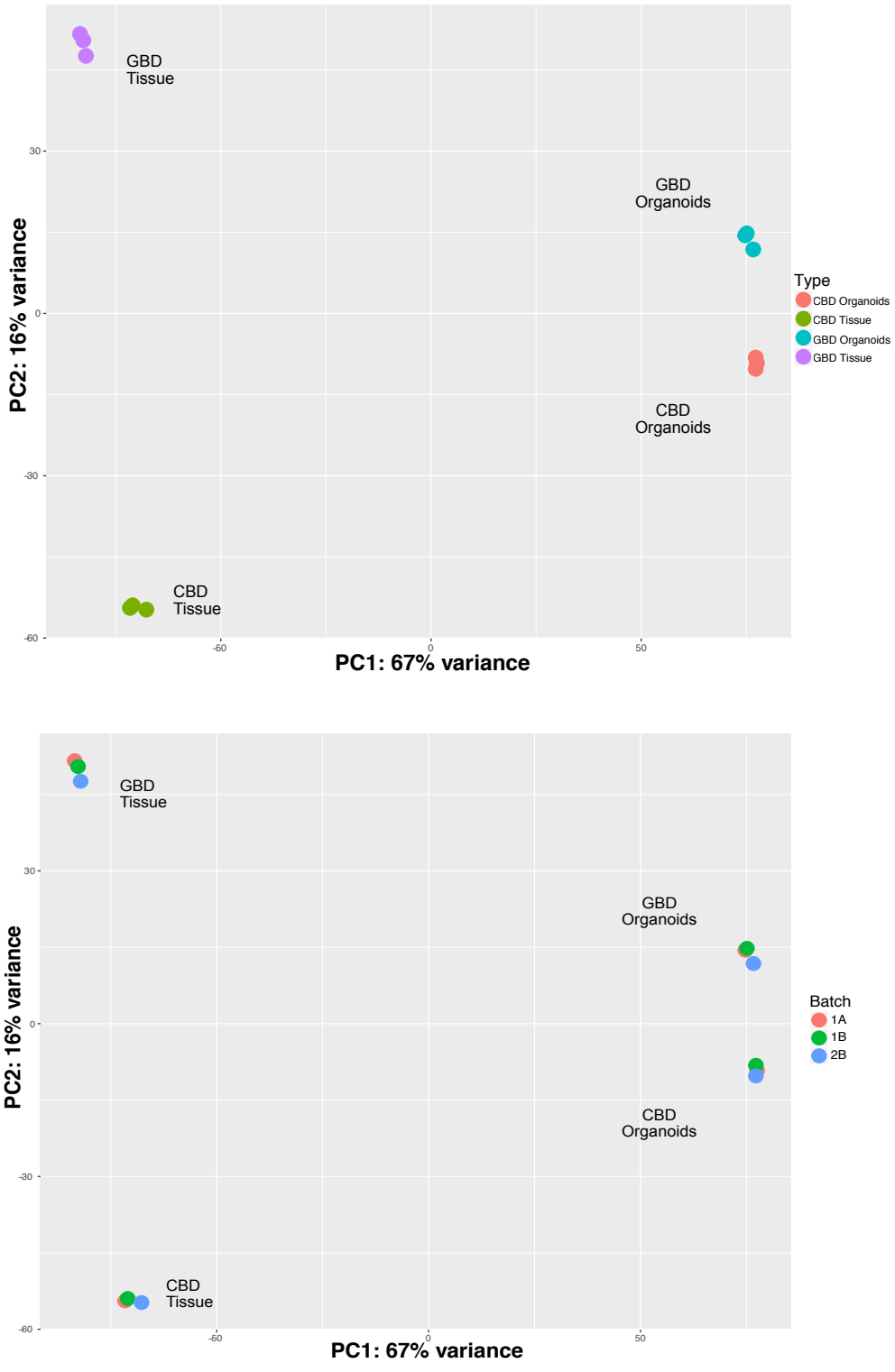


Figure 4.3 Principal component analysis of samples sequenced in all three batches.

Principal component analysis of the *DeSeq2* variance stabilized counts for all genes (19,376) in samples which were sequenced in all three batches labelled by type or by batch.

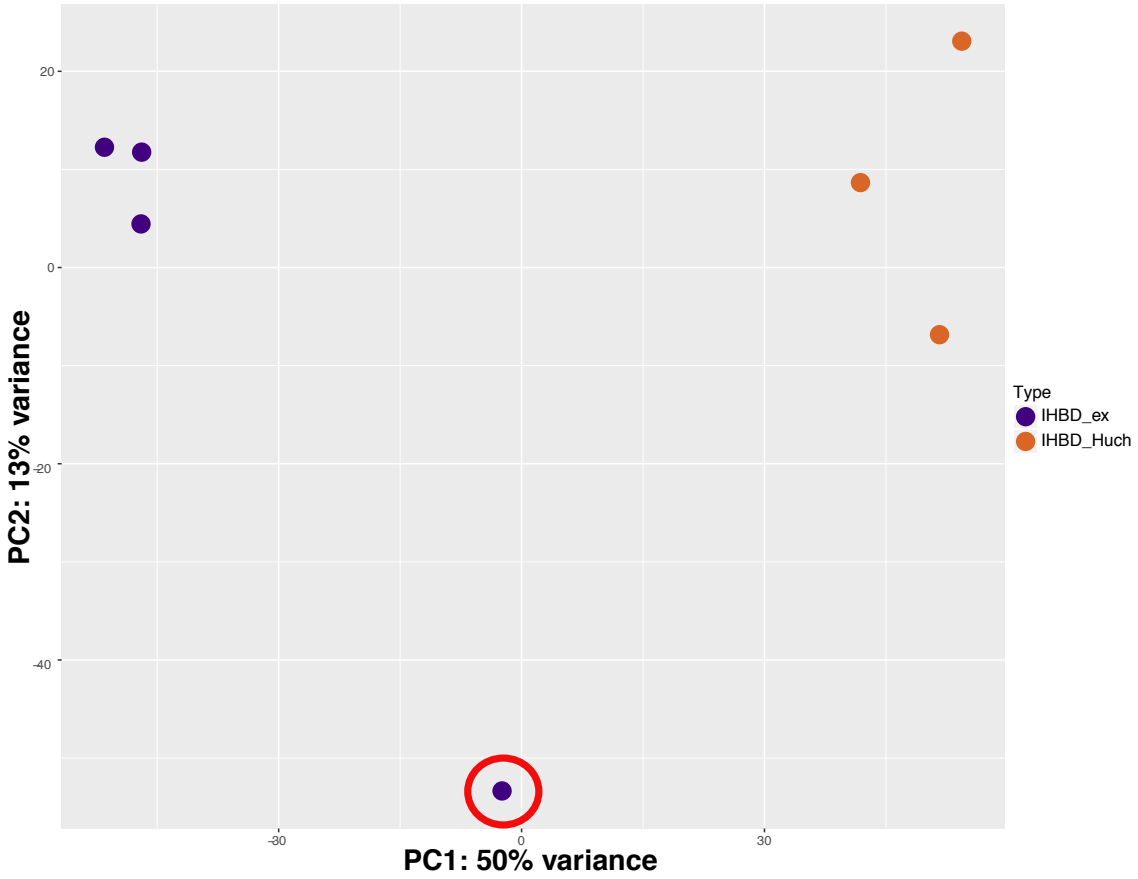


Figure 4.4 Principal component analysis of IHBD organoids to assess for outliers.

Principal component analysis of the *DeSeq2* variance stabilized counts for all genes (19,376) in IHBD_ex and IHBD_Huch organoids demonstrates one of the IHBD_ex samples is an outlier (red circle).

We also noted that IHBD tissue samples clustered very far from the EHBD tissue samples by Principal Component Analysis (PCA) (**Figure 4.2**). The IHBD tissue was obtained from previously frozen liver tissue that had been thawed, enzymatically dissociated, and the bile ducts manually picked, making contamination with other cell types likely. We included these samples as a preliminary exploratory attempt to compare IHBD and EHBD tissues. Given the known limitations of the IHBD tissue samples, we excluded these samples from our initial main analyses. However, these samples will be discussed in section 4.3.8 in the context of preliminary and exploratory analyses.

4.3.3 Transcriptomic profiling of extrahepatic biliary tissues reveals distinct signatures between anatomic regions

To assess what differences may exist between the organoids and their regions of origin, we decided to establish a transcriptional profile for the CBD, GBD, and PancD tissues. Of note, the epithelial enriched tissue samples used for these analyses were collected as described for organoid derivation without any purification step in order to avoid cellular stress associated with dissociation and sorting. Thus, we cannot exclude a limited contamination by non-biliary cell types in the samples. Accordingly, hormonal markers were detected in PancD tissue (i.e. PPY, SST), thereby a low level of contaminating pancreas tissue could explain this gene signature. Interestingly, it has been shown in both human and rodent studies that there may be populations of hormone producing cells which reside in the pancreatic ductal epithelial system and this gene signature could alternatively represent a true population of cells in the pancreatic duct.[70,72,141] Nonetheless, expression of biliary markers such as KRT7, KRT19, SOX9 and EPCAM could be detected at a similar level in the three tissues confirming enrichment for ductal epithelial cells (**Figure 4.5**).

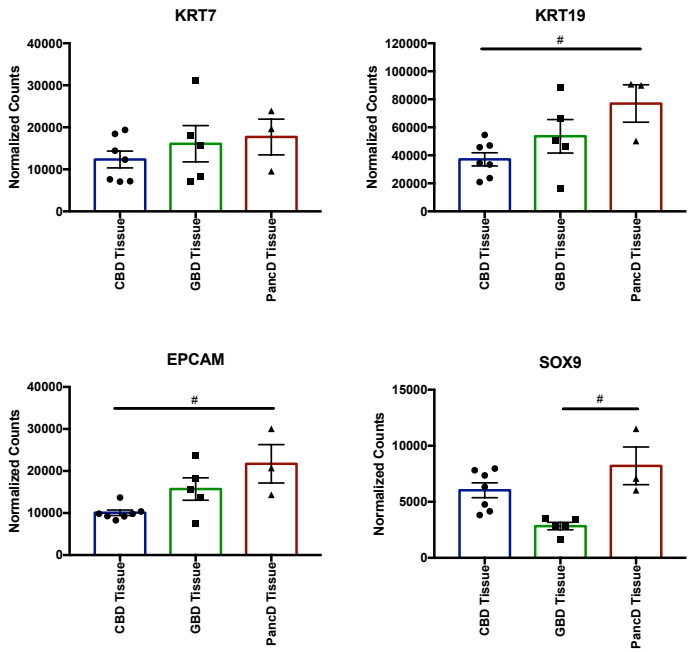


Figure 4.5 Normalized counts of biliary epithelial specific markers for the three extrahepatic tissue regions.

DeSeq2 normalized counts for four bile duct epithelial specific markers. # = false discovery rate (adj p-value) < 0.05 and fold change > 1.

Based on this reassuring observation, we performed PCA, which revealed clustering of samples by tissue type as did correlation analysis (**Figure 4.6 and 4.7**).

Further, differential gene expression (DGE) analyses demonstrated a number of genes were differentially expressed between the tissue regions at a cut-off of false discovery rate (adj p-value) < 0.05 and absolute log_2FoldChange > 1 (**Figure 4.8, Annex File 1**).

3,102 genes were differentially expressed between GBD and PancD tissues with 1,665 genes upregulated in GBD tissue and 1,437 genes upregulated in PancD tissue. 3,715 genes were differentially expressed between CBD and PancD tissues with 1,884 genes upregulated in CBD tissue and 1,831 genes upregulated in PancD Tissues. 1,481 genes were differentially expressed between CBD and GBD tissues with 716 genes upregulated in CBD tissue and 765 genes upregulated in GBD tissue.

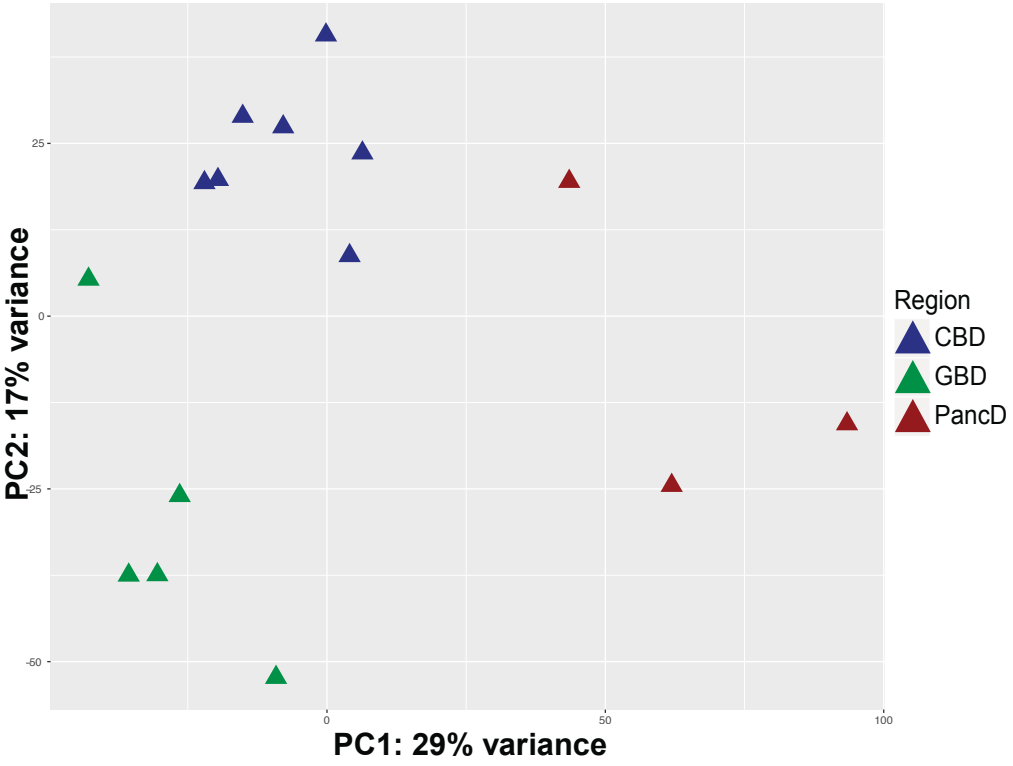


Figure 4.6 Principal component analysis of extrahepatic bile duct tissues

Principal component analysis of the *DeSeq2* variance stabilized counts for the top 5,000 most variable genes between Common Bile Duct (CBD, n=7), Gallbladder (GBD n=5), and Pancreatic Duct (PancD, n=3) tissues.

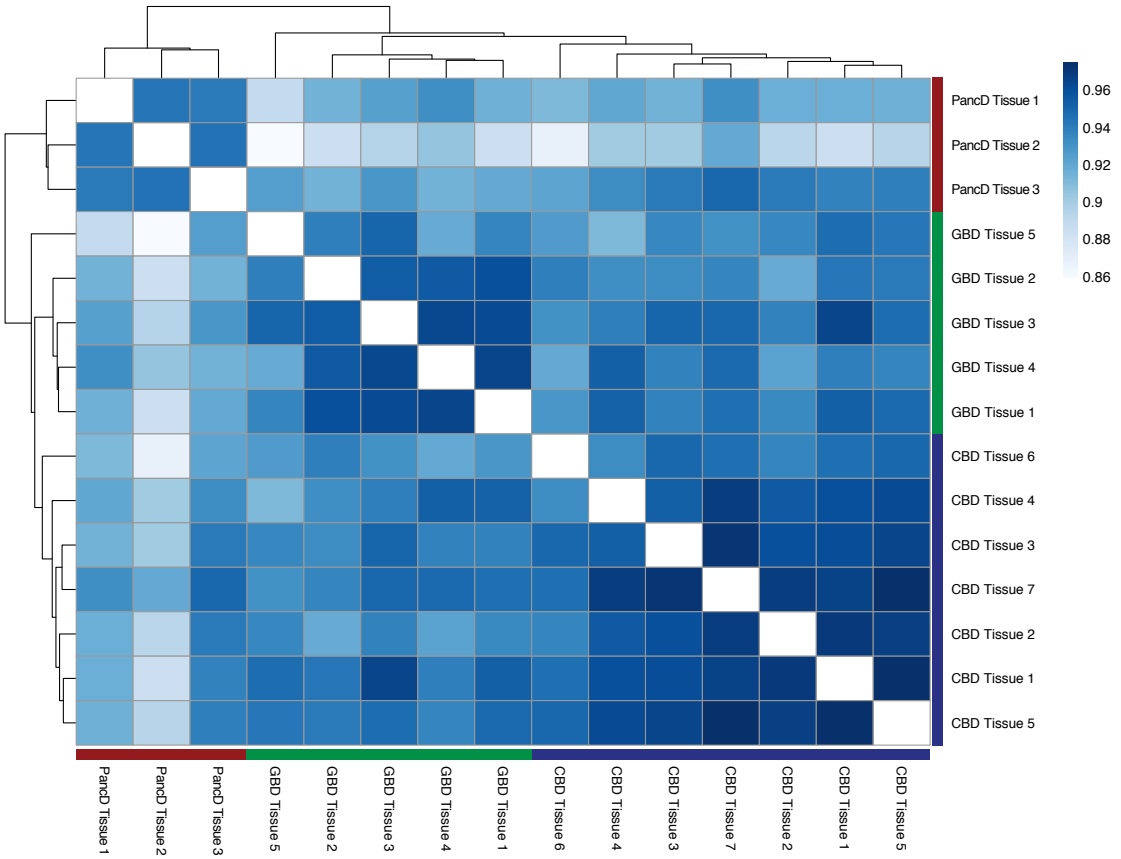


Figure 4.7 Pearson correlation matrix of extrahepatic bile duct tissues

Pearson correlation matrix and clustering for pancreatic duct, gallbladder, and common bile duct tissue samples. Matrix was generated using the R *pheatmap* package.

Using these gene lists, we identified genes likely to be tissue-specific as those which are upregulated in a given tissue type compared to both of the other two tissues. 419 genes were determined to be GBD tissue-specific, 967 PancD tissue-specific, and 256 CBD specific (**Figure 4.8, Annex File 2**). Using this set of genes, we performed hierarchical clustering analysis and showed that both the samples and genes cluster by tissue region (**Figure 4.9**).

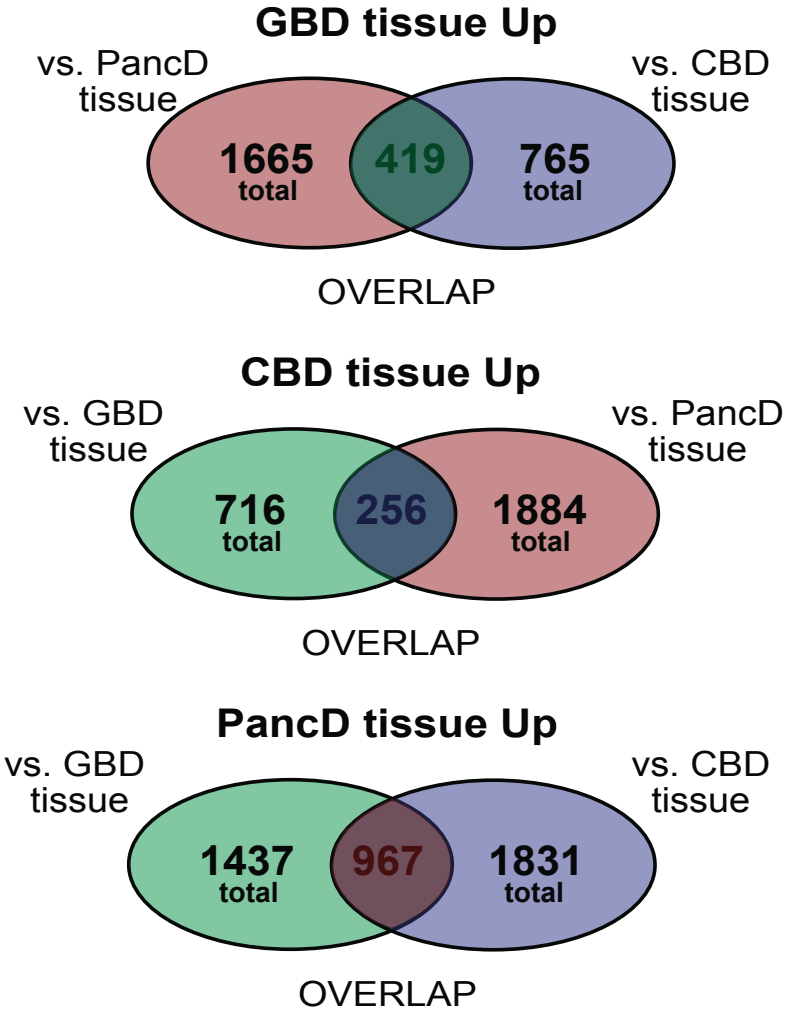


Figure 4.8 Number of genes differentially expressed between the three extrahepatic bile duct tissue regions.

Venn diagrams showing the number, as well as overlap, of genes upregulated in either GBD, CBD, or PancD tissues in comparison to each of the other two tissues. Significance cut-off: abs(log_2Fold Change) >1 and adj p-value < 0.05.

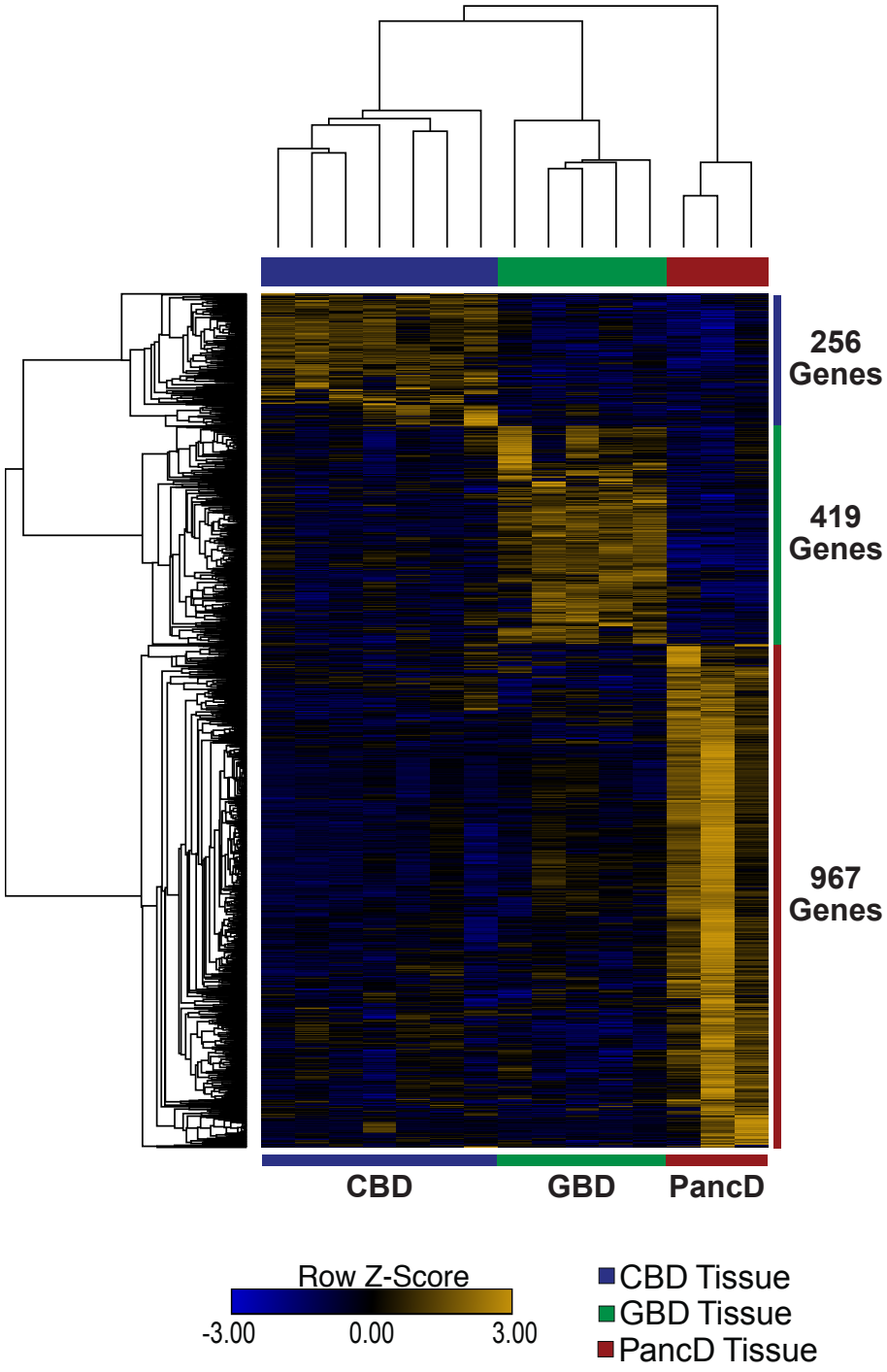


Figure 4.9 Heatmap and hierarchical clustering analysis of genes specifically upregulated in extrahepatic tissue region.

Heatmap of the tissue-specific genes upregulated in CBD compared to both GBD and PancD tissues (n=256 genes), GBD compared to both PancD and CBD tissues (n=419 genes), and PancD compared to both CBD and GBD tissues (n=967 genes). Hierarchical clustering performed on the variance stabilized transformed counts using one minus Pearson correlation distance with average linkage.

Gene ontology (GO) analyses on the tissue-specific gene lists and the genes differentially expressed between pairs of tissues yielded interesting insight into functional differences, which may exist between the regions. The 419 GBD specific genes included ones such as UGT1A6, MOGAT1, CA4, SOX17, VDR and were enriched for GO terms (adj p-value < 0.05) involved in xenobiotic, lipid, carbohydrate, and steroid metabolic processes, among others (**Figure 4.10, Annex File 3).** Similar enrichment was observed when comparing genes upregulated in GBD tissue to either CBD or PancD tissues separately (**Figure 4.11 and 4.12, Annex File 3**), suggesting that of these EHBD regions, GBD tissue plays the largest role in active metabolic processes.

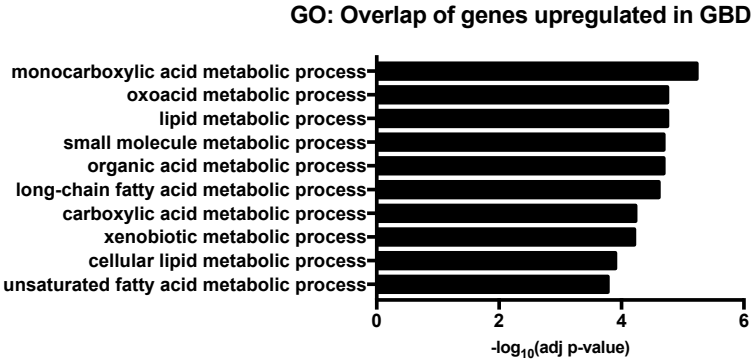


Figure 4.10 Top ten enriched gene ontology terms for 419 genes upregulated in GBD tissue compared to both CBD and PancD tissues.

Top ten most significant gene ontologies (biological processes) for 419 genes upregulated in GBD tissue compared to both CBD and PancD.

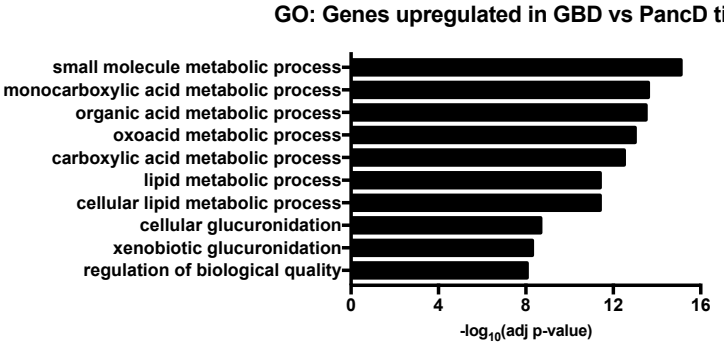


Figure 4.11 Top ten enriched gene ontology terms for genes upregulated in GBD tissue compared to PancD tissue.

Top ten most significant gene ontologies (biological processes) for 1,665 genes upregulated in GBD compared to PancD tissue.

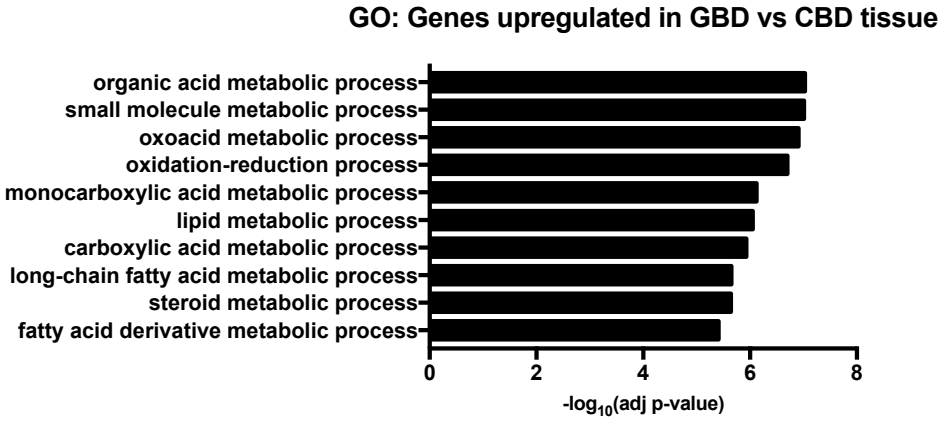


Figure 4.12 Top ten enriched gene ontology terms for genes upregulated in GBD tissue compared to CBD tissue.

Top ten most significant gene ontologies (biological processes) for 765 genes upregulated in GBD compared to CBD tissue.

The 967 PancD specific genes included ones such as SCTR and CDX2 and were enriched for GO terms involved in protein targeting and localization to the membrane or endoplasmic reticulum, among others (**Figure 4.13**, **Annex File 3**). Similar GO terms were observed when comparing genes upregulated in PancD tissue compared with GBD or CBD tissue alone (**Figure 4.14 and Figure 4.15, Annex File 3**).

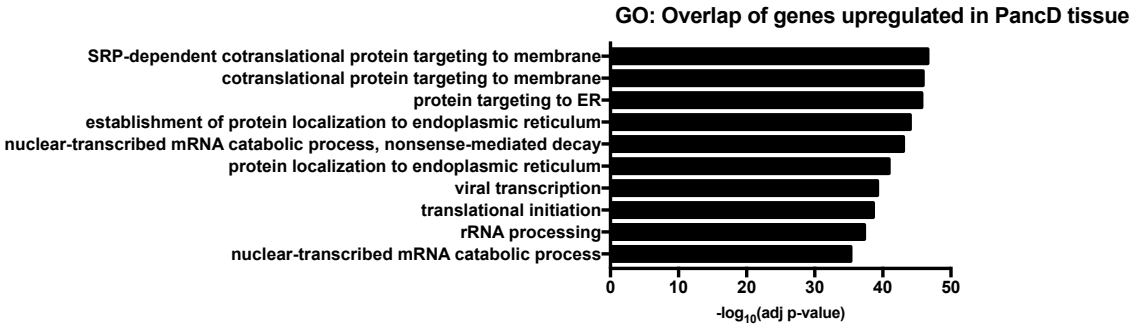


Figure 4.13 Top ten enriched gene ontology terms for 967 genes upregulated in PancD tissue compared to both GBD and CBD tissues.

Top ten most significant gene ontologies (biological processes) for 967 genes upregulated in PancD compared to GBD and CBD tissues.

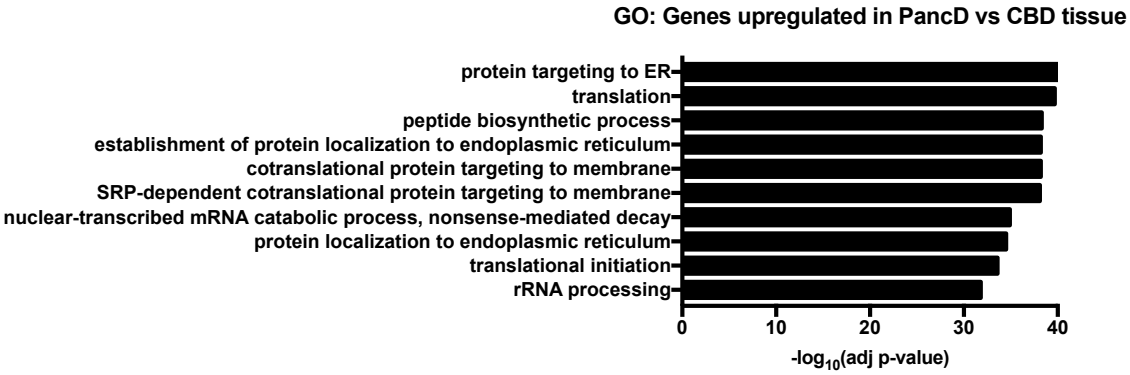


Figure 4.14 Top ten enriched gene ontology terms for genes upregulated in PancD compared to CBD tissue.

Top ten most significant gene ontologies (biological processes) for 1831 genes upregulated in PancD compared to CBD tissue.

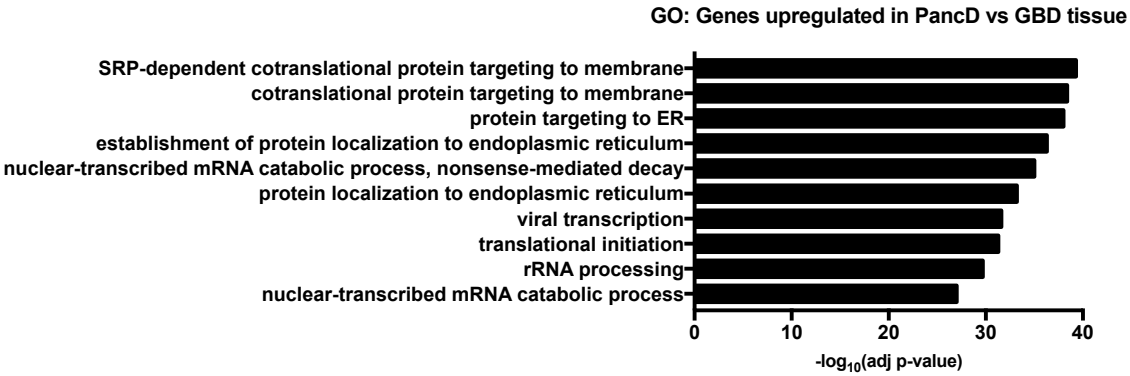


Figure 4.15 Top ten enriched gene ontology terms for genes upregulated in PancD compared to GBD tissue.

Top ten most significant gene ontologies (biological processes) for 1437 genes upregulated in PancD compared to GBD tissue.

The 256 CBD specific genes included ones such as HOXB2, HOXB3, and ABCA1 but were only enriched significantly for five GO terms with tissue morphogenesis and signal transduction among them (**Figure 4.16, Annex File 3**). Genes upregulated in CBD compared to GBD tissue were enriched for GO terms such as developmental processes, tube development, and tube morphogenesis (**Figure 4.17, Annex File 3**). Genes upregulated in CBD tissue compared to PancD tissue were enriched for GO terms such as xenobiotic glucuronidation and lipid transport (**Figure 4.18, Annex File 3**).

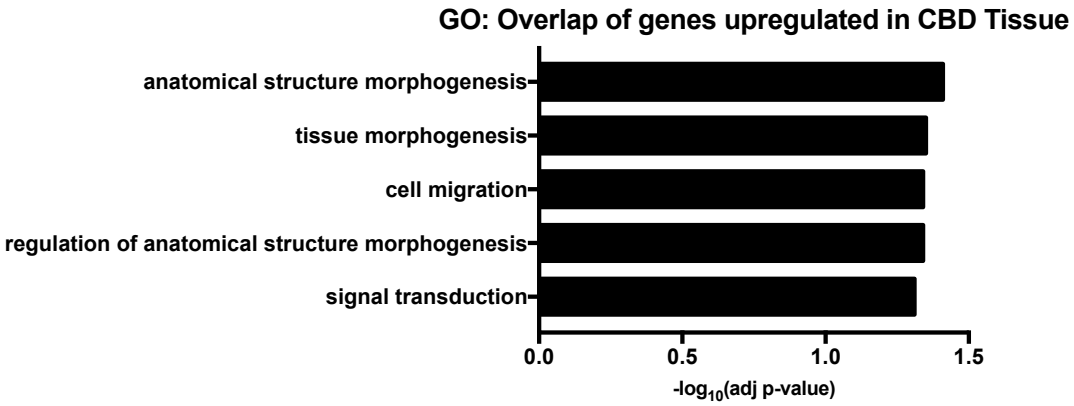


Figure 4.16 Five significantly enriched gene ontology terms for 256 genes upregulated in CBD compared to both GBD and PancD tissues.

Five gene ontology terms (biological processes) were significantly enriched for the 256 genes upregulated in CBD compared to both GBD and PancD tissues.

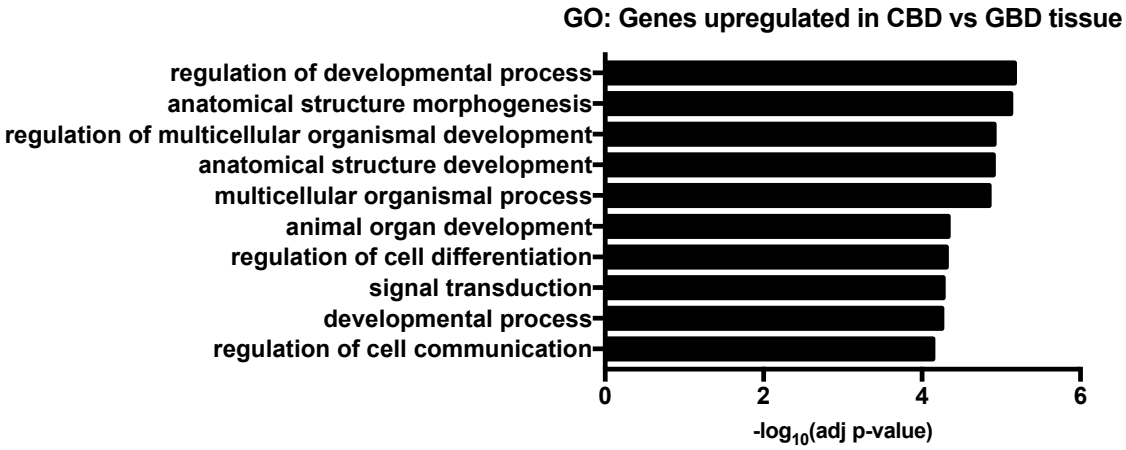


Figure 4.17 Top ten significantly enriched gene ontology terms for genes upregulated in CBD compared to GBD tissue.

Top ten most significant gene ontology terms (biological processes) for 716 genes upregulated in CBD compared to GBD tissue.

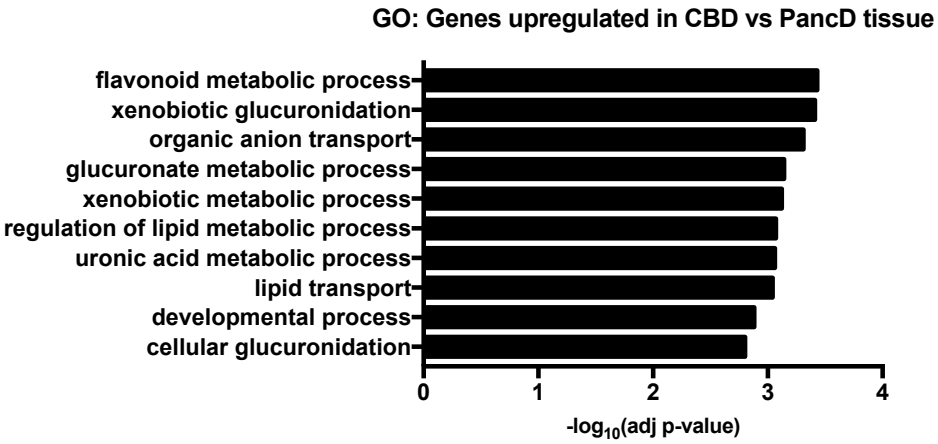


Figure 4.18 Top ten significantly enriched gene ontology terms for genes upregulated in CBD compared to PancD tissue.

Top ten most significant gene ontology terms (biological processes) for 1884 genes upregulated in CBD compared to PancD tissue.

We also looked at the genes that were upregulated in both GBD and CBD tissue in comparison to PancD tissue, and found 856 overlapping genes (**Figure 4.19, Annex File 4**). This list included genes known to be involved in lipid, cholesterol, bile acid, and xenobiotic/bilirubin related metabolic processes (i.e. FGF19, NR1I2, VDR, UGT1A4, UGT1A6, UGT1A10, ABCA1, APOA1, PPARG), which are known constituents of bile secretions that are only in contact with GBD/CBD epithelium and not the PancD (**Figure 4.20**). GO analyses further confirmed this observation (**Figure 4.21, Annex File 5**).

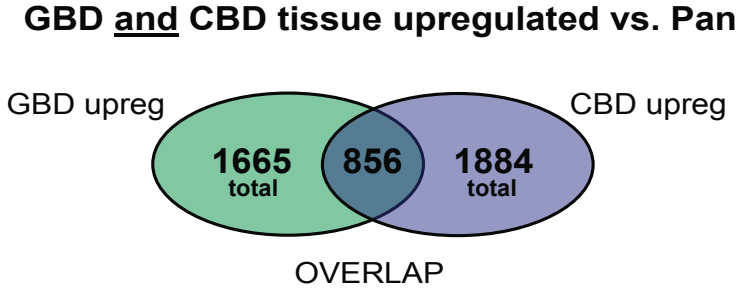


Figure 4.19 Overlap of genes differentially expressed between GBD and CBD tissues compared to PancD tissue.

Venn diagram showing the number, as well as overlap, of genes upregulated in either GBD compared to PancD tissue or CBD compared to PancD tissue. Significance cut-off: abs (log_2Fold Change) >1 and adj p-value < 0.05.

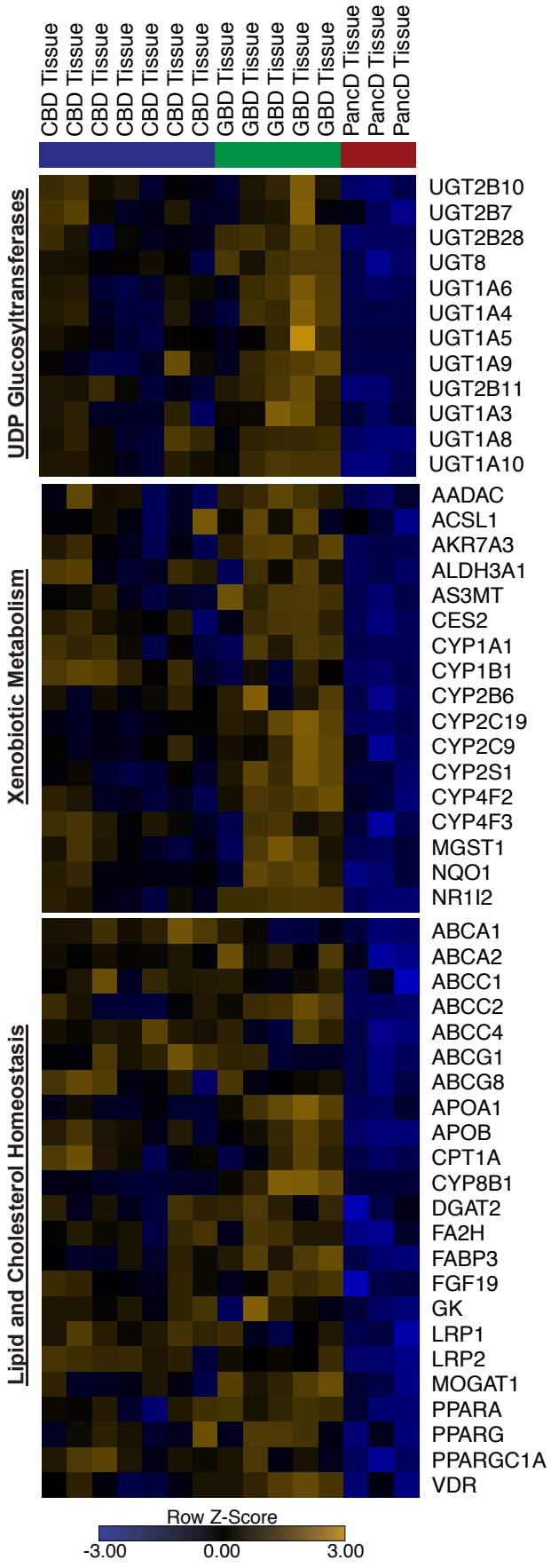


Figure 4.20 Heatmap of selected genes involved in lipid, cholesterol, bile acid, and bilirubin related metabolic processes upregulated in GBD and CBD tissues compared to PancD tissues.

Heatmap of the variance stabilized counts of selected genes that are upregulated in both GBD and CBD compared to PancD tissues involved in lipid, cholesterol, xenobiotic related metabolic processes.

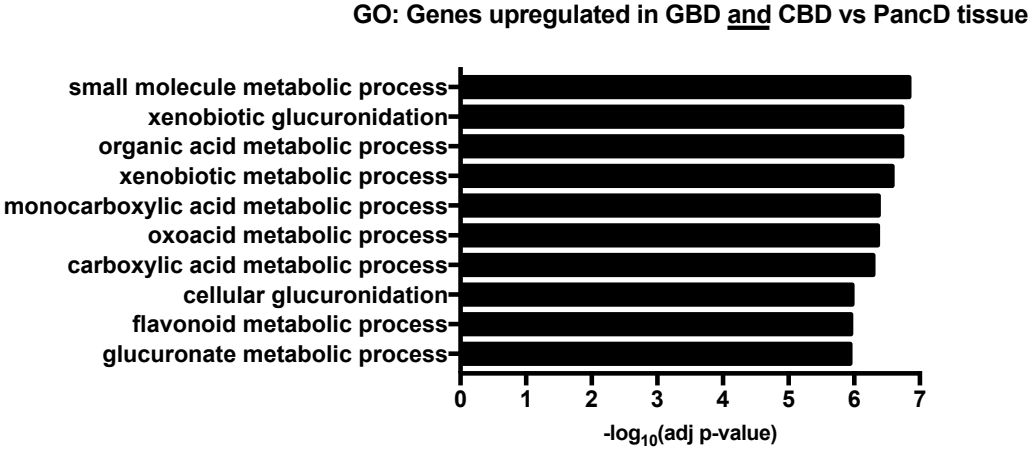


Figure 4.21 Top ten significantly enriched gene ontology terms for genes upregulated in both GBD and CBD tissues compared to PancD tissue.

Top ten most significant gene ontology terms (biological processes) for 856 genes upregulated in both GBD and CBD tissues compared to PancD tissue.

It is also interesting to note that of the 856 genes upregulated in both GBD and CBD, compared to PancD tissue, 108 of these were additionally upregulated in GBD compared to CBD tissue. These 108 genes were enriched for GO terms similar to those for the 856 genes (**Figure 4.22, Annex File 5**). Taken together, this supports the conclusion that anatomical niche shapes the expression profile of cholangiocytes within the biliary tree, with the GBD appearing to play the largest role in xenobiotic, lipid, and cholesterol metabolism.

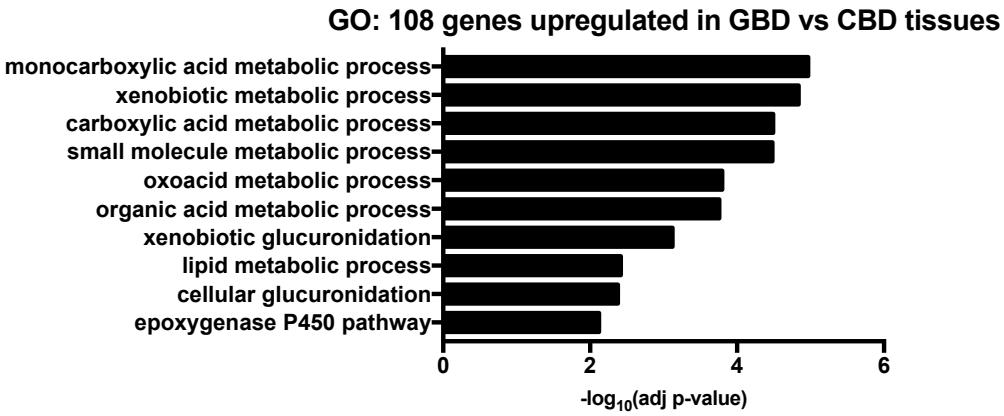


Figure 4.22 Top ten most significantly enriched gene ontology terms for 108 genes upregulated in GBD tissue compared to CBD tissue that were also upregulated in CBD tissue in comparison to PancD tissue.

Top ten most significant gene ontology (biological processes) terms. Out of 856 genes that were upregulated in both GBD and CBD in comparison to PancD tissues, 108 genes were also found to be upregulated in GBD compared to CBD tissue.

4.3.4 Comparing extrahepatic bile duct tissue RNA-Sequencing results to previously published works

As part of the Human Protein Atlas, Kampf et al (2014) performed RNA-sequencing on GBD tissue from three donors and compared GBD tissue to 27 other human tissues in order to determine GBD enriched genes.[138] To further validate and benchmark our own data, we compared our data to this published work. Kampf et al (2014) previously identified 140 genes with increased expression in human GBD tissue in comparison to 27 other tissue types, but this did not include comparisons to CBD or PancD tissues. Of these genes, we found 131 represented in our dataset. The nine genes not included in our dataset represented either genes no longer annotated in the current ENSEMBL database or non-protein coding genes, which were not included in our analyses.

Five genes were identified as being GBD specific with expression levels five-fold higher in GBD tissue than in any other tissue. We first looked to see whether these five genes were represented in our dataset and whether they remained GBD tissue-specific even when compared to more closely related tissues such as CBD and PancD. Four of the genes: FGF19, CHST4, MOGAT1, and UGT2B28 were expressed in our dataset. The fifth gene, AC083862.1 has been re-classified as a pseudogene in the current ENSEMBL

annotation and was therefore not included in our dataset of protein-coding genes. MOGAT1 and UGT2B28 were found to be upregulated in GBD tissue in comparison to both PancD and CBD tissues (**Figure 4.23**). Further, FGF19 and CHST4 were upregulated in GBD and CBD, compared to PancD tissues, but were not differentially expressed between GBD and CBD tissues.

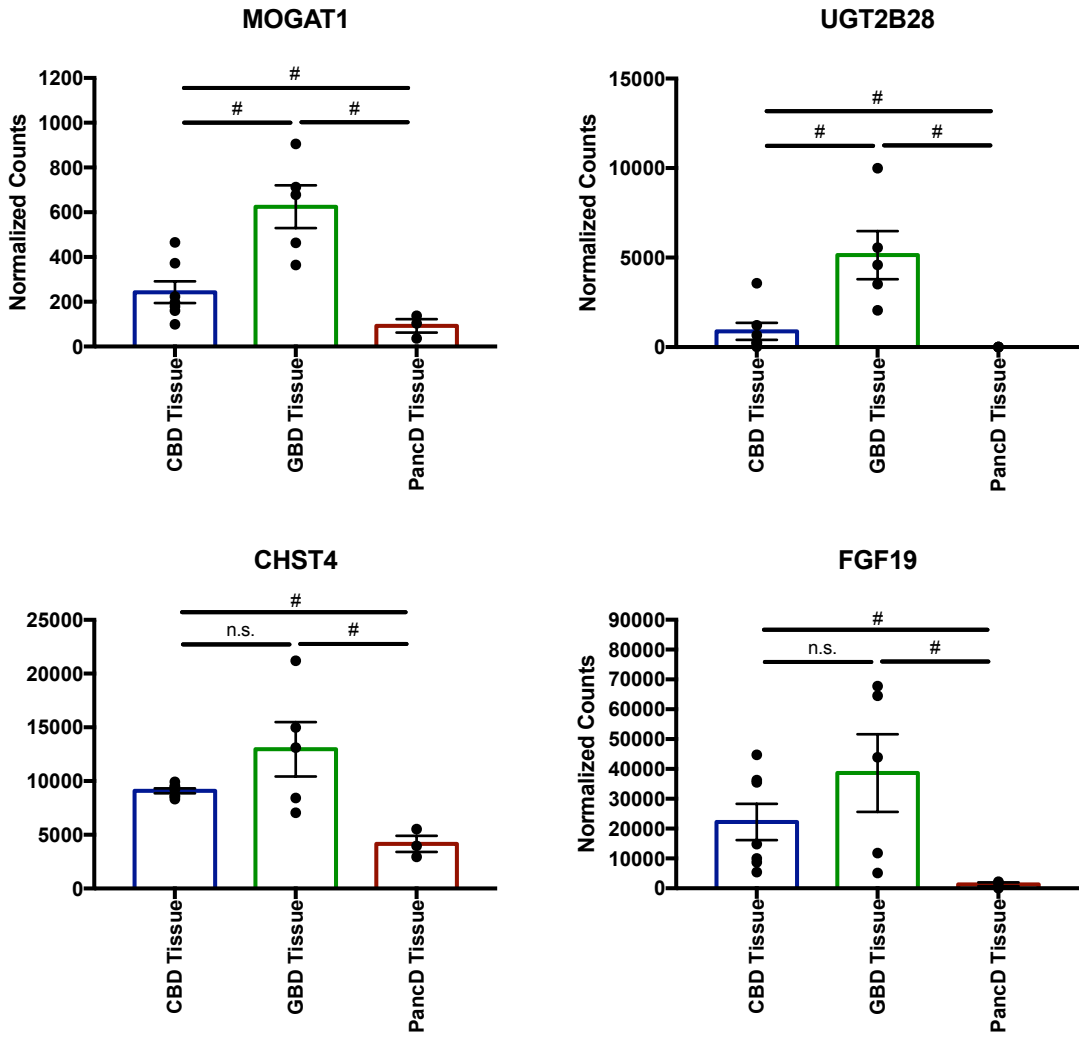


Figure 4.23 Normalized counts for four genes previously identified as being Gallbladder-tissue specific by Kampf et al (2014)

DeSeq2 normalized counts for the three extrahepatic bile duct tissue regions for four genes previously identified as being gallbladder-tissue specific when comparing to 27 other human tissue types. # = false discovery rate (adj p-value) < 0.05 between the groups and fold change > 1. n.s. = not significant

We further compared our data with the list of 131 gallbladder-enriched genes by first comparing them to our list of 419 genes which were differentially expressed in GBD tissue compared to both CBD and PancD tissues. We found fifteen genes overlapped (**Figure 4.24**). A further ten genes were found to overlap the list of 856 genes which had been found to be upregulated in both GBD and CBD tissue compared to PancD tissue. Interestingly, twelve genes, which were GBD enriched in the work by Kampf et al (2014), were found to be upregulated in either PancD tissue or CBD tissue when compared to GBD tissue in our dataset. This suggests that some genes, although having enrichment in GBD tissue, may actually be enriched in the other two extrahepatic tissue types to a greater degree. Lastly, 78 genes from the Kampf et al (2014) list were not differentially expressed between any of the three tissues in our dataset. These genes included ones such as CFTR and ONECUT2, which are ubiquitously expressed in biliary epithelial cells. Therefore, these 78 genes likely represent basal cholangiocyte specific genes, which are not differentially expressed between the three extrahepatic tissue regions, but are simply enriched in biliary tissues compared to other non-related tissue types (**Figure 4.25**).

By comparing our data with that of Kampf et al (2014) we were able to validate our own dataset by confirming that certain genes, such as MOGAT1 and UGT2B28, are gallbladder-tissue enriched even when compared to CBD and PancD tissues. Furthermore, it was found that the gallbladder enriched genes CHST4 and FGF19 are not specific to GBD tissue but are also expressed in CBD tissues, but are not highly expressed in PancD tissue. Ultimately, this continues to support our results by suggesting differences in the transcriptional profile exist for these three tissue regions.

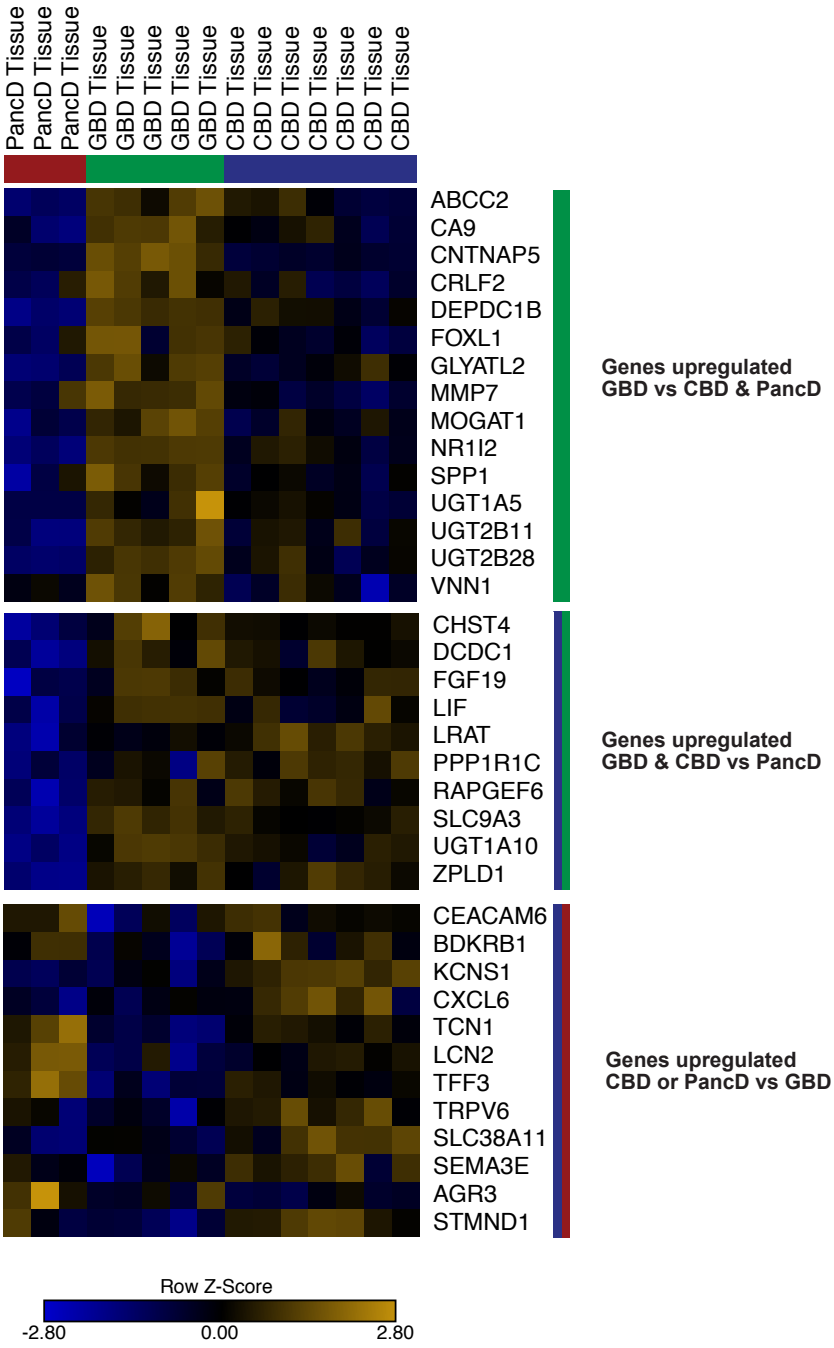


Figure 4.24 Heatmap of gallbladder tissue enriched genes previously identified by Kampf et al (2014) that were identified in our dataset.

Heatmap of the variance stabilized counts for the three extrahepatic bile duct tissue regions. Genes shown were ones identified as being in common with our dataset and that of Kampf et al (2014). The first cluster shows 15 genes upregulated in GBD compared to both CBD and PancD tissues which were also in the list of gallbladder-enriched genes identified by Kampf et al (2014). The second cluster contains 10 genes upregulated in GBD and CBD compared to PancD tissues that were also in the gallbladder-enriched list. The third cluster contains 12 genes upregulated in either CBD or PancD compared to GBD tissue despite being in the gallbladder-enriched gene list.

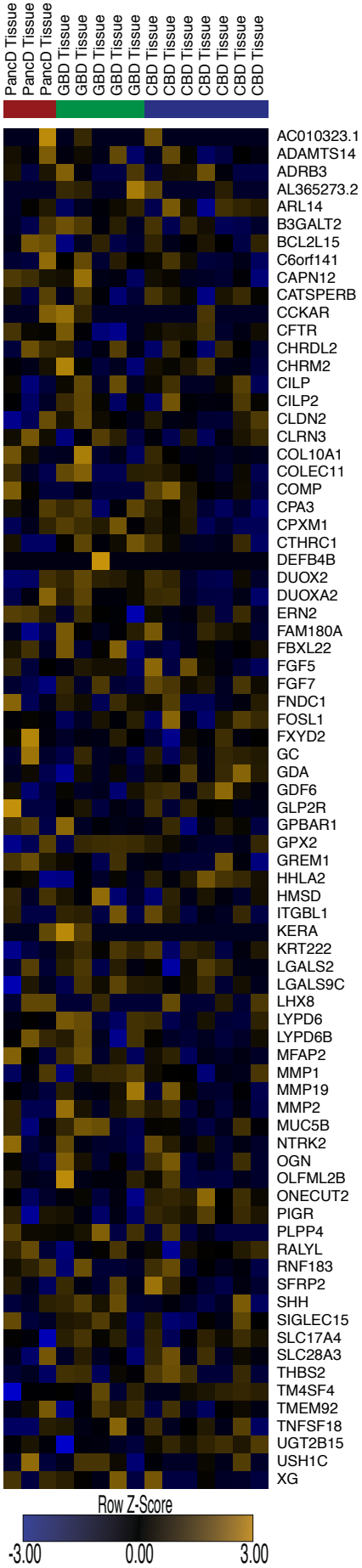


Figure 4.25 Heatmap of genes identified by Kampf et al (2014) as enriched in gallbladder tissue that are not differentially expressed between GBD, CBD, or PancD tissues

Heatmap of the variance stabilized transformed counts for genes not differentially expressed between GBD, PancD, or CBD tissues that were previously identified by Kampf et al (2014) as being enriched in GBD tissues.

4.3.5 Transcriptomic profiling of extrahepatic biliary organoids reveals downregulation of mature biliary markers and upregulation of cell cycle and WNT signalling genes

In other epithelial derived organoid systems, such as the intestine, organoids maintain a remarkable similarity to their tissue of origin and express markers of both adult stem cells and differentiated cell lineages.[124] As biliary tissue is made up of predominantly one mature cell type, the cholangiocyte, we wanted to also investigate how extrahepatic biliary organoids compare to their tissue of origin and whether any regional tissue-specific markers are maintained. To do so, RNA-sequencing was performed on organoids derived from CBD, GBD, and PancD tissues. PCA revealed that the largest source of variation in the samples (56%) was between tissues and organoids and that tissue-specific differences only accounted for 10% of the variation between samples (**Figure 4.26**). The three different organoid types clustered closely. However, clustering by tissue of origin was still discernible.

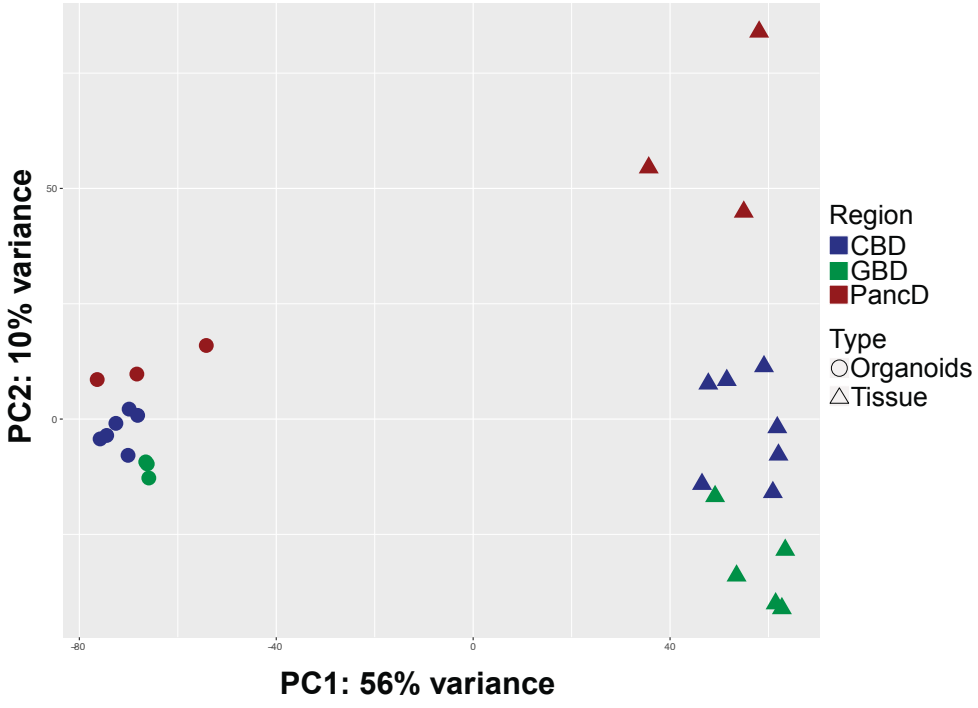


Figure 4.26 Principal component analysis of extrahepatic bile duct tissues and organoids

Principal component analysis of variance stabilized counts for the top 5,000 most variable genes between Common Bile Duct (CBD) tissue (n=7), CBD organoids (n=6), Gallbladder (GBD) tissue (n=5), GBD organoids (n=3), Pancreatic Duct (PancD) tissue (n=3), and PancD organoids (n=3).

In line with the large separation between tissue and organoid samples by PCA, DGE analyses showed a large number of genes were differentially expressed between each organoid type and their corresponding tissue of origin (5,586 genes, CBD tissue vs organoids; 4,537 genes, GBD tissue vs organoids; and 5,594 genes, PancD tissue vs organoids) (**Figure 4.27, Annex File 6**). To explore these differences, we examined the expression levels of selected genes known to be involved in mature biliary epithelial function, adult stem cells, WNT signalling, and cell cycle regulation which were differentially expressed between organoids and tissues (**Figure 4.28)**.

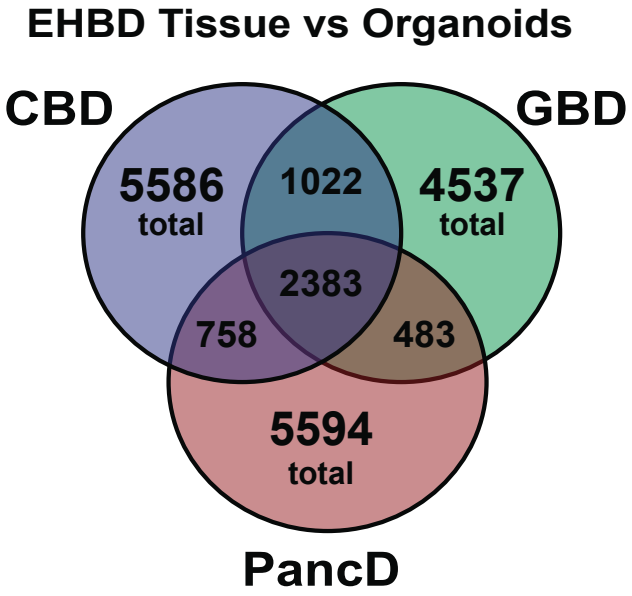


Figure 4.27 Number of genes differentially expressed between the extrahepatic bile duct tissues and corresponding organoids.

Venn diagram showing the number, and overlap, of genes differentially expressed between CBD tissue and organoids, PancD tissue and organoids, or GBD tissue and organoids. Significance cut-off: log_2Fold Change >1 and adj p-value < 0.05.

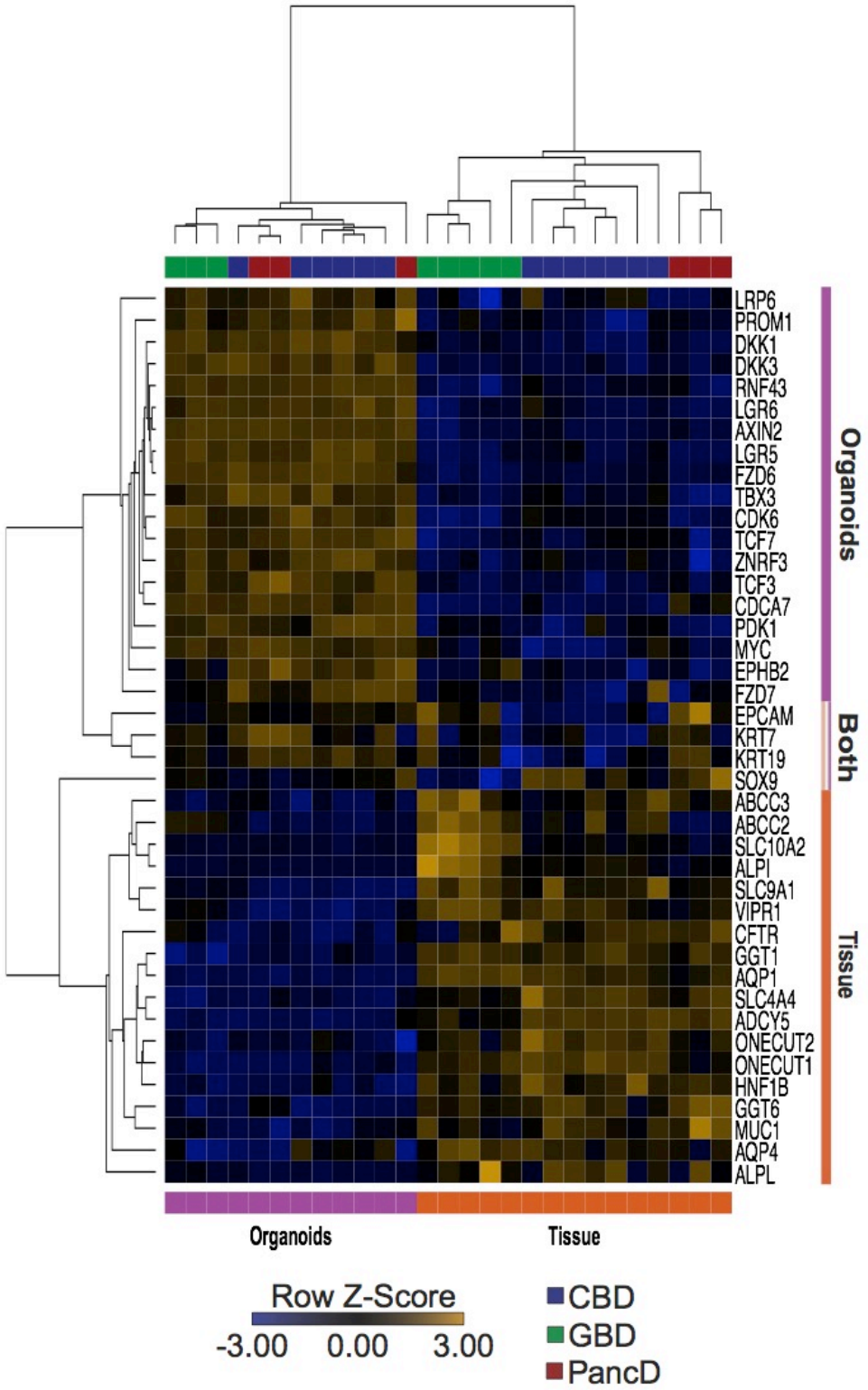


Figure 4.28 Heatmap and hierarchical clustering analysis of selected biliary and adult stem cell genes comparing EHBD organoids and primary tissues.

Heatmap of selected biliary and adult stem cell/WNT signalling genes. Hierarchical clustering performed on the variance stabilized transformed counts using one minus Pearson correlation distance with average linkage.

Hierarchical clustering of these genes highlighted the differences between organoids and primary tissues, with organoids having significantly lower expression of functional biliary markers (i.e. GGT1/6, CFTR, MUC1, VIPR1) and significantly higher expression of adult stem cell markers and genes known to be involved in WNT and cell cycle pathways (i.e. LGR5/6, PROM1, AXIN2, CDK6). Ductal markers such as KRT7, KRT19, EPCAM, and SOX9 were not differentially expressed between tissues and organoids suggesting that these genes are intrinsic biliary markers, which are not affected by WNT or *in vitro* culture conditions. Further, GO analyses of the top 200 differentially expressed genes (genes with the smallest adj p-value) between each of the organoids and their corresponding primary tissues were enriched for GO terms involved in cell cycle processes and cell division. Genes upregulated in tissues compared to organoids were enriched for GO terms involved in secretion, transport, and metabolic processes (**Figure 4.29, Annex File 7**).

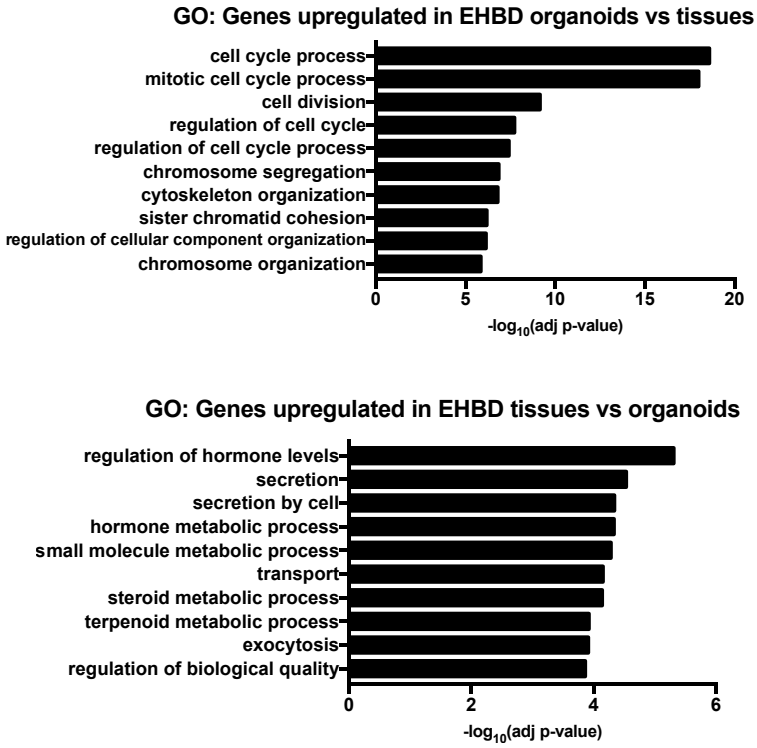


Figure 4.29 Top ten significantly enriched gene ontology terms for the top 200 differentially expressed genes in EHBD organoids compared to primary tissues.

Top ten most significant gene ontologies (biological processes) for the top 200 differentially expressed genes upregulated in extrahepatic tissues compared to organoids or upregulated in extrahepatic organoids compared to tissues.

We also assessed whether any regional tissue-specific signatures were maintained in the EHBD organoids. It was found that, on average, only 45% of significantly differentially expressed genes between CBD, GBD, and PancD derived organoids overlapped with differentially expressed genes previously identified in their corresponding primary tissues (**Figure 4.30, Annex File 8 and 9**). 91.6% of these genes showed the same directionality of expression when comparing organoid and tissue expression patterns. This suggests that EHBD organoids, despite differing from their tissue of origin, do maintain some regional specific markers in culture.

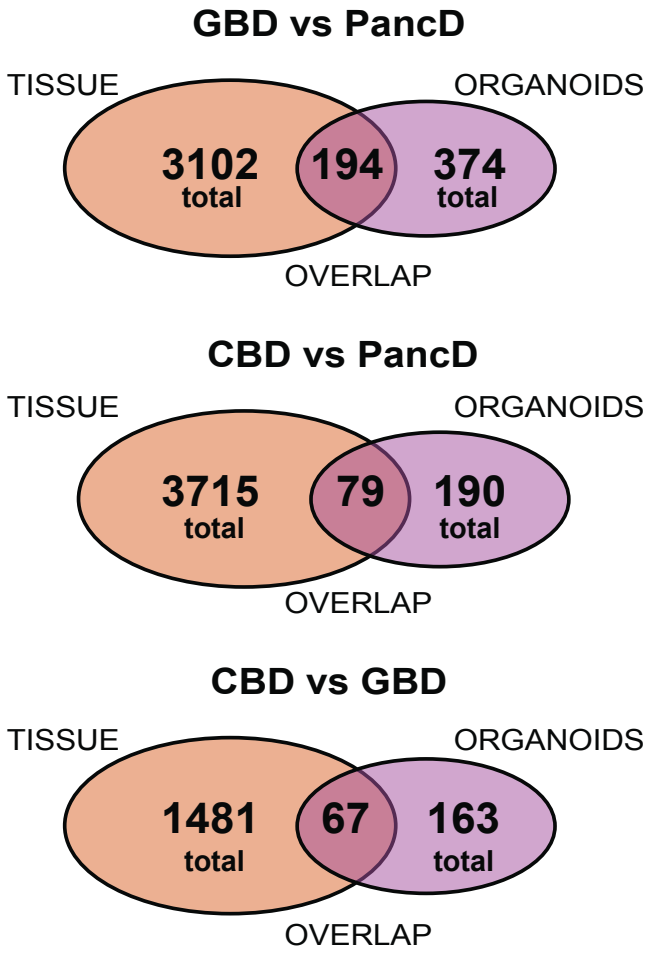


Figure 4.30 Number of genes differentially expressed between the extrahepatic bile duct tissues, between the organoids, and the number of genes overlapping.

Venn diagrams displaying number, as well as overlap, of genes differentially expressed comparing GBD versus PancD tissues and *in vitro* organoids, CBD versus PancD tissues and *in vitro* organoids, and CBD versus GBD tissues and *in vitro* organoids. Significance cut-off: abs (log_2Fold Change) >1 and adj p-value < 0.05

We validated the expression of three of these genes, SOX17, CDX2, HOXB2, by qPCR (**Figure 4.31**) and also of SOX17 and CDX2 at the protein-level by IF (**Figure 4.32**). Unfortunately, we were unable to find an antibody against HOXB2 that worked on frozen tissue sections. However, our RNA-Seq data and also qPCR validation of HOXB2 suggests that it is a CBD specific marker.

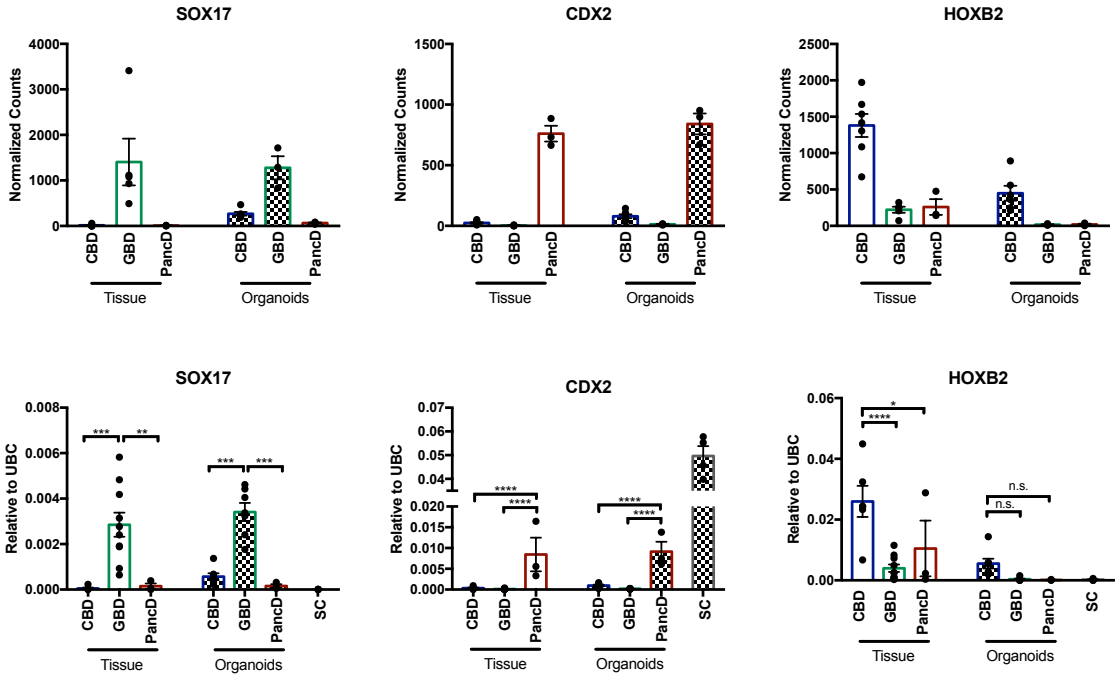


Figure 4.31 RNA-Seq normalized read counts and qPCR validation of three tissue specific markers which are maintained in the organoids.

RNA-Seq read counts and qPCR validation of three genes found to retain tissue-specific expression patterns in organoid cultures. For qPCR validation: CBD Tissue (n=6), CBD Organoids (n=7), GBD Tissue (n=10), GBD Organoids (n=7), PancD Tissue (n=3), and PancD organoids (n=3) and gene expression is normalized to the housekeeping gene UBC and data is plotted as mean and SEM. *=$p \leq 0.05$, **=$p \leq 0.01$, ***=$p \leq 0.001$, ****$p \leq 0.0001$, n.s. = not significant. Sigmoid colon (SC) organoids (n=4) were used as controls.

SOX17 was only expressed in GBD tissue and showed the highest expression level in GBD organoids. CBD organoids, despite having lower expression of SOX17 at the RNA level, did show expression by IF. PancD organoids were negative for SOX17. CDX2 was exclusively expressed in PancD tissue and organoids. These expression patterns did not change significantly over time in culture (**Figure 4.33**).

Taken together, these data show that biliary epithelial cells lose, in part, their regional identity when grown as organoids in the presence of WNT signalling. Nonetheless, this seems to only be partial as the expression of a few, but very specific, regional markers are maintained in EHBD organoids.

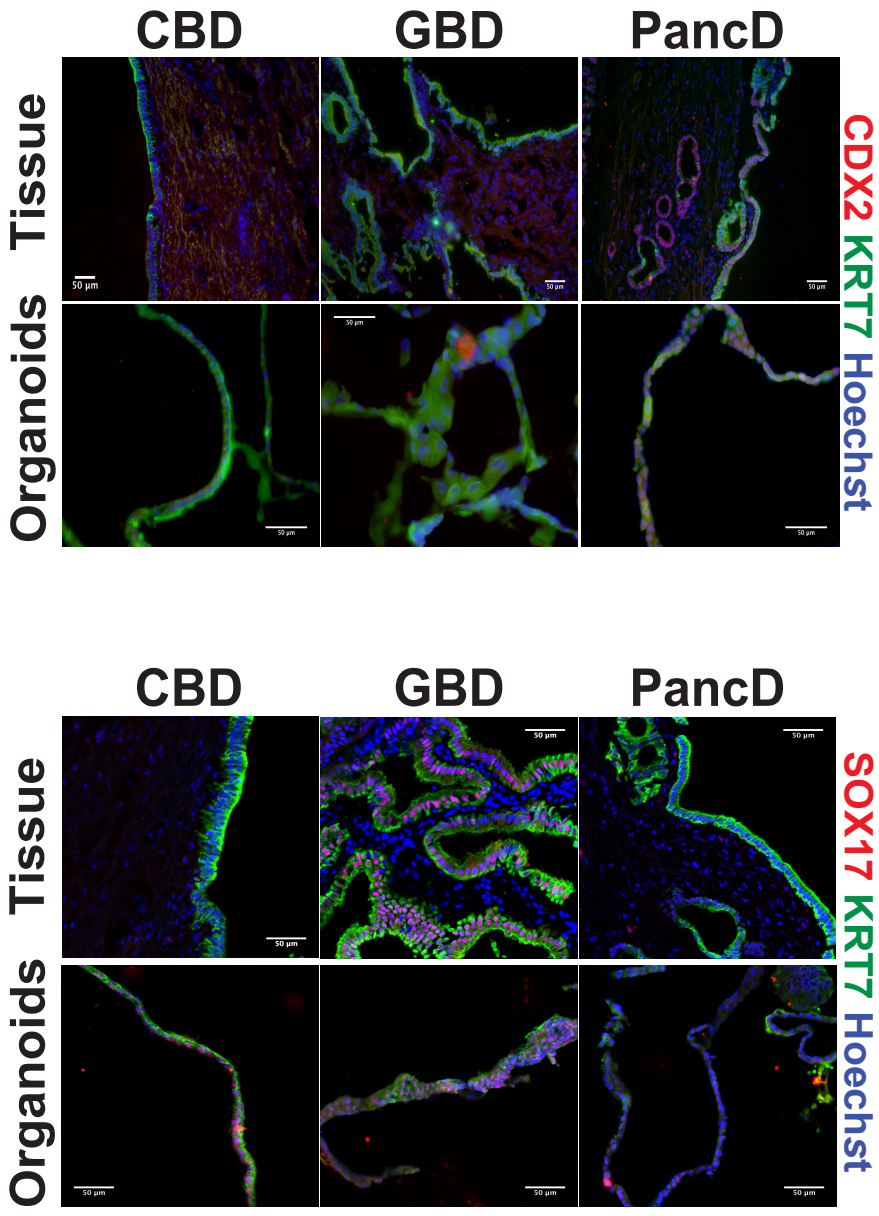


Figure 4.32 Immunofluorescence analysis of EHBD organoids and primary tissues for two genes found to be differentially expressed between tissue regions.

Scale bar=50 μm

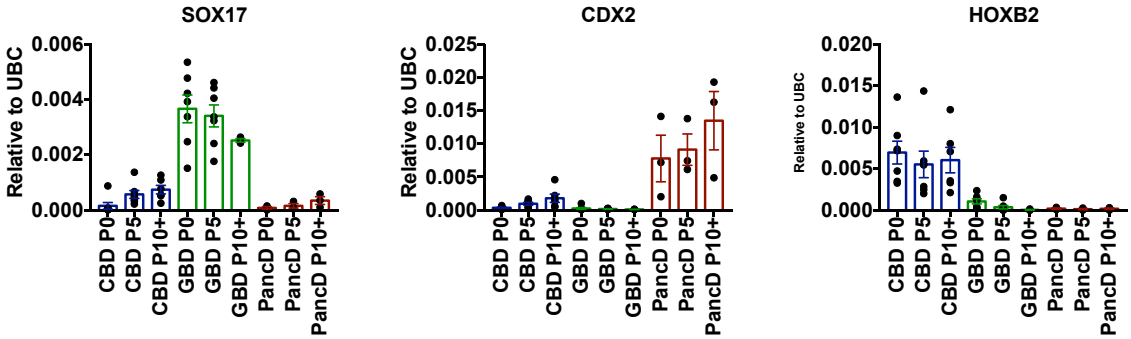


Figure 4.33 Expression of SOX17, CDX2, and HOXB2 over time in culture in extrahepatic bile duct organoids.

qPCR analyses showing the expression of SOX17, CDX2, and HOXB2 in extrahepatic bile duct organoids (n=3-7, n=2-7 donor lines per passage) from common bile duct (CBD), gallbladder (GBD), and pancreatic duct (PancD) over time in culture (P0, P5, P10 or higher). Gene expression is normalized to the housekeeping gene UBC and data is plotted as mean and SEM.

4.3.6 Comparison of extrahepatic bile duct organoids to intestinal LGR5+ stem cell signature from previously published works.

In comparing our EHBD organoids to their tissues of origin, it was found that many functional biliary markers were down-regulated and genes such as LGR5/6 and PROM1, which are commonly associated with adSCs, were upregulated. To explore this potential stem cell signature further, we compared our EHBD organoids to a previously published gene signature for LGR5+ intestinal stem cells described in work performed by Munoz et al (2012).[142] The work by Munoz et al (2012) was performed in mouse and utilized the transgenic *LGR5-EGFP-ires-CreERT2* knock-in mouse which labels LGR5 positive cells with GFP, allowing them to be selectively sorted from intestinal tissue and only the LGR5 positive cells to be analysed. Munoz et al (2012) performed both microarray profiling and also proteomic analyses using mass spectrometry to identify genes that were enriched in LGR5+ intestinal stem cells compared to non-LGR5+ intestinal cells from freshly sorted cells. In doing so, they found 510 genes that were enriched in LGR5+ cells and deemed these genes to make up an *intestinal stem cell signature*. Unfortunately, as there are currently no reliable LGR5 antibodies that stain human tissue, a human intestinal stem cell signature has not yet been described.

However, many of the genes found by Munoz et al (2012) have been shown to be markers of human intestinal stem cells grown in organoid cultures suggesting this *intestinal stem*

cell signature discovered in the mouse may be informative also for human adult stem cells.[88]

Given this, we first took the list of 510 *intestinal stem cell signature* genes and determined whether human orthologs exist for the genes and if the genes were represented in our own RNA-Sequencing data set. Of the 510 genes, we found 493 genes which met these criteria (**Annex File 10**). Next, genes which were upregulated in EHBD organoids compared to their tissue of origin were compared to this list of 493 genes. It was found that a number of genes overlapped between the lists (**Figure 4.34, Annex File 10**). This suggested that EHBD organoids, when compared to their primary tissues, appear to upregulate some of these *intestinal stem cell signature* genes. Further, a large number of these genes were in common across the three different organoid types, with 135 of the intestinal stem cell signature genes being upregulated in at least two out of the three EHBD organoid types in comparison to their primary tissues. A heatmap of those 135 genes is shown in **Figure 4.35**.

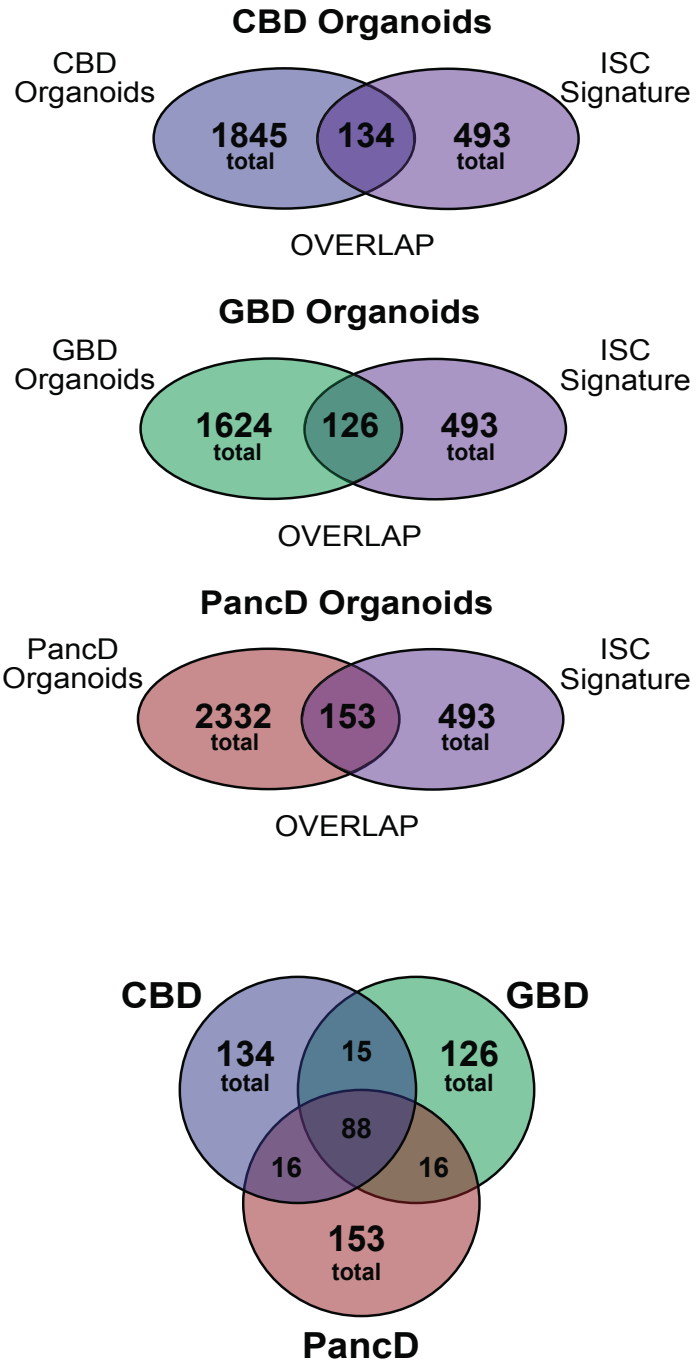


Figure 4.34 Genes upregulated in EHBD organoids compared to their tissues of origin which overlap with an intestinal stem cell signature of 493 genes described by Munoz et al (2012).

The top three diagrams show the number of genes that were upregulated in extrahepatic bile duct organoids in comparison to their tissues of origin (1,845 common bile duct, 1,624 gallbladder, and 2,332 pancreatic duct organoids) and the number of these genes that overlap a 493 gene signature reported by Munoz et al (2012) for intestinal LGR5+ adult stem cells (134 common bile duct, 126 gallbladder, 153 pancreatic duct). The bottom diagram shows the overlap of these overlapping genes between the three different organoid types.

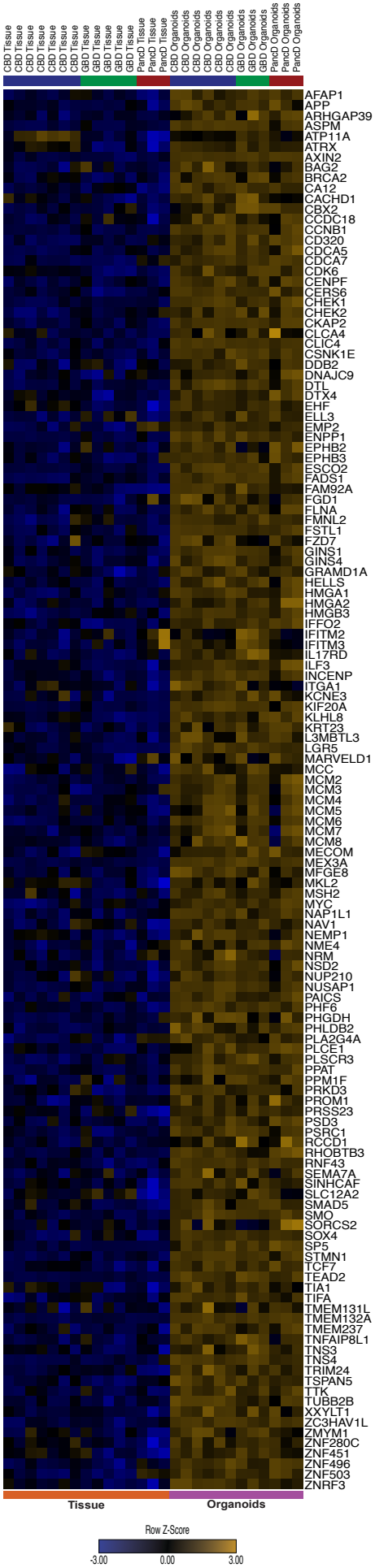


Figure 4.35 Intestinal stem cell signature genes described by Munoz et al (2012) that are upregulated in extrahepatic bile duct organoids compared to their tissues of origin.

Heatmap of the variance stabilized counts depicting the 135 *intestinal stem cell signature* genes which were upregulated in at least two out of the three EHBD organoid types in comparison to their primary tissues.

We also performed GO enrichment analyses on these 135 genes and found they were largely involved in cell cycle processes and WNT signalling (**Figure 4.36, Annex File 10**). This comparison further supports our earlier conclusions that EHBD organoids differ greatly from their tissue of origin and upregulate genes involved in cell cycle progression and WNT signalling. It is interesting to find that many of these genes are also over-represented in a population of cells known to be well-established as adult stem cells in the intestine.

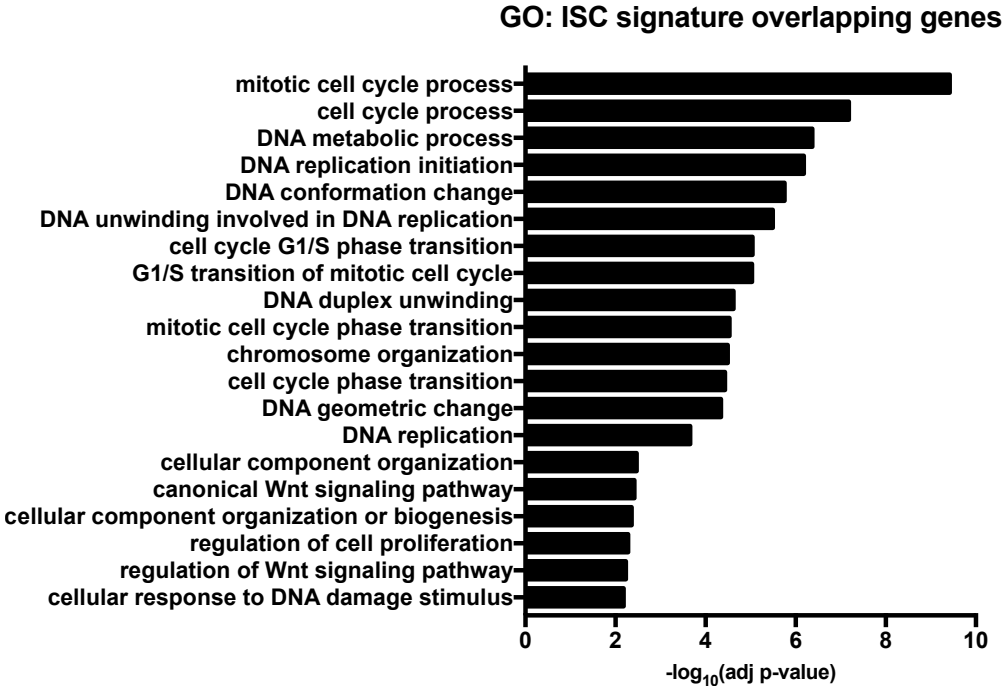


Figure 4.36 Gene ontology analysis of the 135 intestinal stem cell signature genes described by Munoz et al (2012) which were upregulated in extrahepatic bile duct organoids in comparison to their tissues of origin.

Top twenty gene ontology terms (biological processes) which were enriched in the 135 intestinal stem cell signature genes described by Munoz et al (2012) which were upregulated in at least two out of the three extrahepatic bile duct organoid types compared to their tissues of origin.

4.3.7 Intrahepatic biliary organoids and extrahepatic biliary organoids display different transcriptional profiles

In addition to performing the basic characterizations of our EHBD and IHBD organoids described in Chapter 3, we also performed sequencing on CBD, GBD, PancD, IHBD_ex, and IHBD_Huch organoids. PCA analysis demonstrated that the largest source of variation was between EHBD organoids and IHBD organoids. EHBD organoids clustered closely together regardless of their tissue of origin, while both IHBD_ex and IHBD_Huch clustered separately from EHBD organoids and each other (**Figure 4.37**). Correlation analysis further confirmed this observation (**Figure 4.38**). As all three of the EHBD organoid types clustered closely and the number of genes differentially expressed between the three organoid types was minimal in our earlier analyses, we combined them into a single "EHBD" group for further analyses.

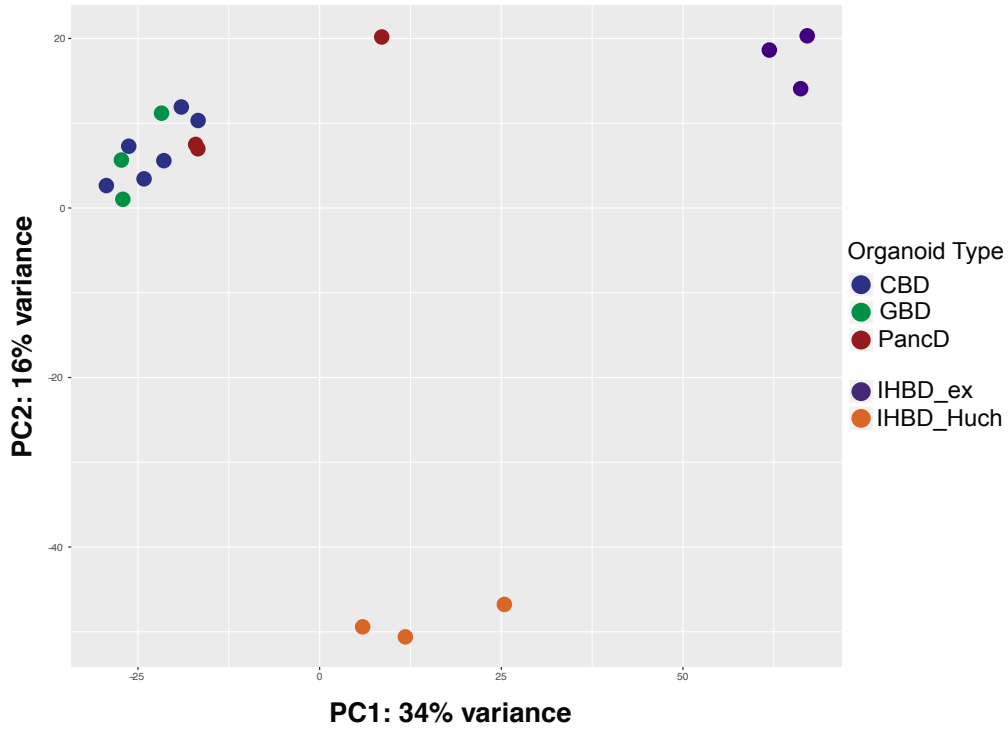


Figure 4.37 Principal component analysis of EHBD and IHBD organoids.

Principal component analysis of variance stabilized counts for the top 5,000 most variable genes between common bile duct (CBD) (n=6), gallbladder (GBD) (n=3), and pancreatic duct (PancD) (n=3) extrahepatic bile duct organoids (EHBD) as well as intrahepatic bile duct (IHBD) organoids cultured in extrahepatic conditions (IHBD_ex, n=3) or in conditions described by Huch et al (2015) (IHBD_Huch (n=3)).

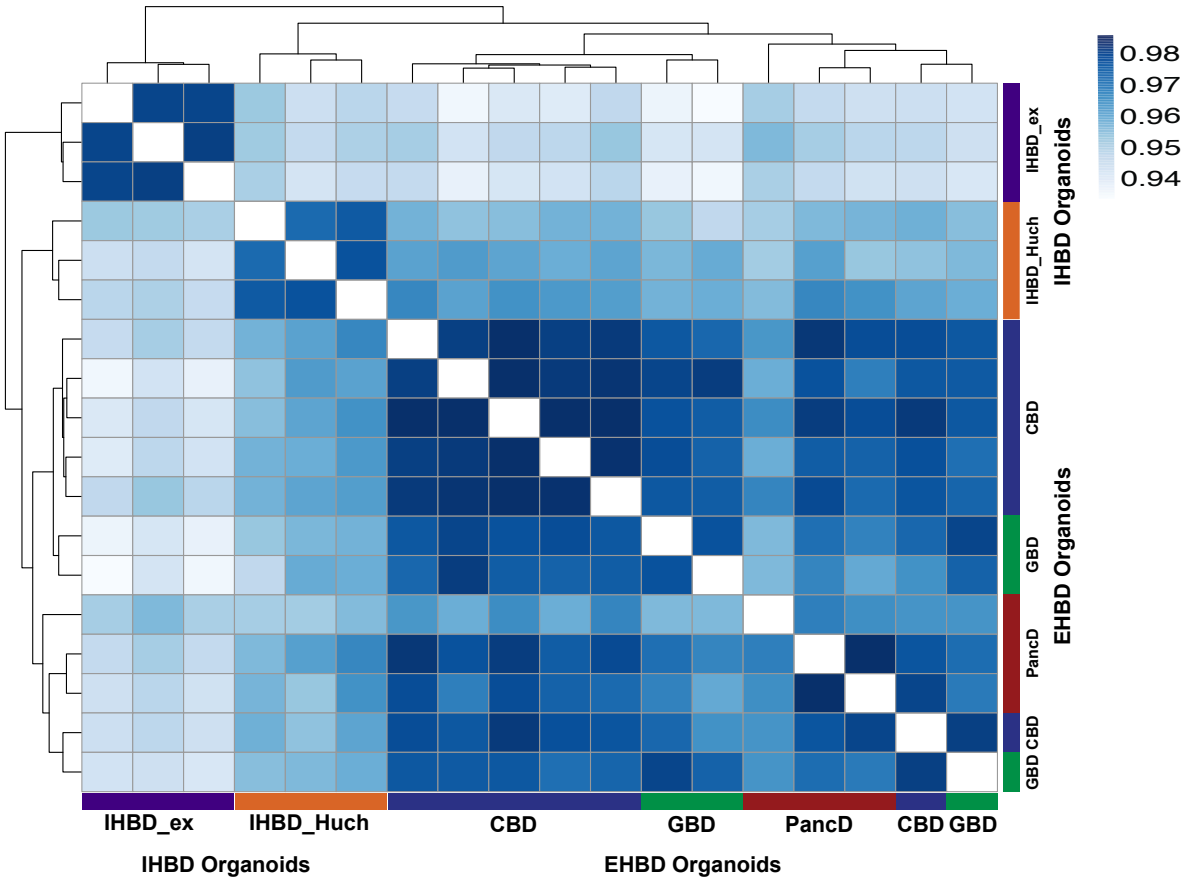


Figure 4.38 Pearson correlation matrix of IHBD and EHBD organoids.

Sample distance matrix of one minus Pearson correlation between EHBD and IHBD organoids. Matrix was generated using the R *pheatmap* package.

We then sought to understand the transcriptional divergence of IHBD_ex organoids from IHBD_Huch and EHBD organoids. 2,122 genes were differentially expressed between IHBD_ex and IHBD_Huch organoids (1,120 upregulated in IHBD_ex and 1,002 upregulated in IHBD_Huch). While 2,990 genes were differentially expressed between IHBD_ex and EHBD organoids (1,411 upregulated in IHBD_ex and 1,579 upregulated in EHBD). Many of these genes overlapped between these comparisons and 612 genes were upregulated in IHBD_ex organoids compared to both IHBD_Huch and EHBD organoids and 544 genes upregulated in IHBD_Huch and EHBD organoids compared to IHBD_ex organoids (**Figure 4.39, Annex File 11**).

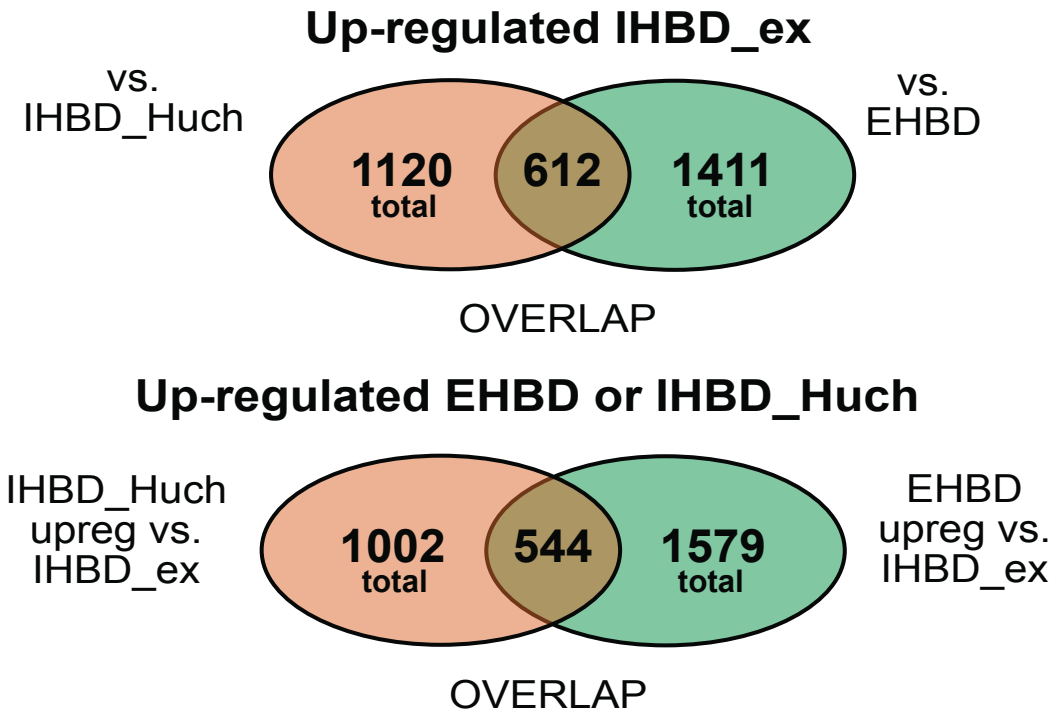


Figure 4.39 Number and overlap of genes differentially expressed between IHBD_ex and EHBD or IHBD_Huch organoids.

Venn diagram showing the number, and overlap, of genes differentially expressed in IHBD_ex organoids compared to IHBD_Huch and EHBD organoids.

As expected, given the proliferation differences observed between the organoids, the 544 genes upregulated in IHBD_Huch and EHBD organoids were enriched significantly for just two GO terms including mitotic cell cycle process (adj p-value = $4.01E10^{-2}$) and cell cycle process (adj p-value = $4.06E10^{-2}$). Interestingly, the 612 genes upregulated in IHBD_ex organoids included genes involved in mature hepatocyte function (i.e. ALB, TTR, NR1H4, CYP7B1, complement/clotting factors) (**Figure 4.40, Annex File 11**). However, the progenitor cell markers (i.e. PROM1, SPP1) were also upregulated, as were markers known to be involved in cell death/senescence and epithelial mesenchymal transition (i.e. CASP1, VIM). Thus, this analysis suggests complex differences between IHBD organoids without uncovering a major source of divergence. Nonetheless, these results clearly demonstrate that EHBD and IHBD organoids are transcriptionally distinct, even when cultured in the same conditions. Further, the extrahepatic culture conditions, and WNT, may promote the senescence and/or differentiation of IHBD cells.

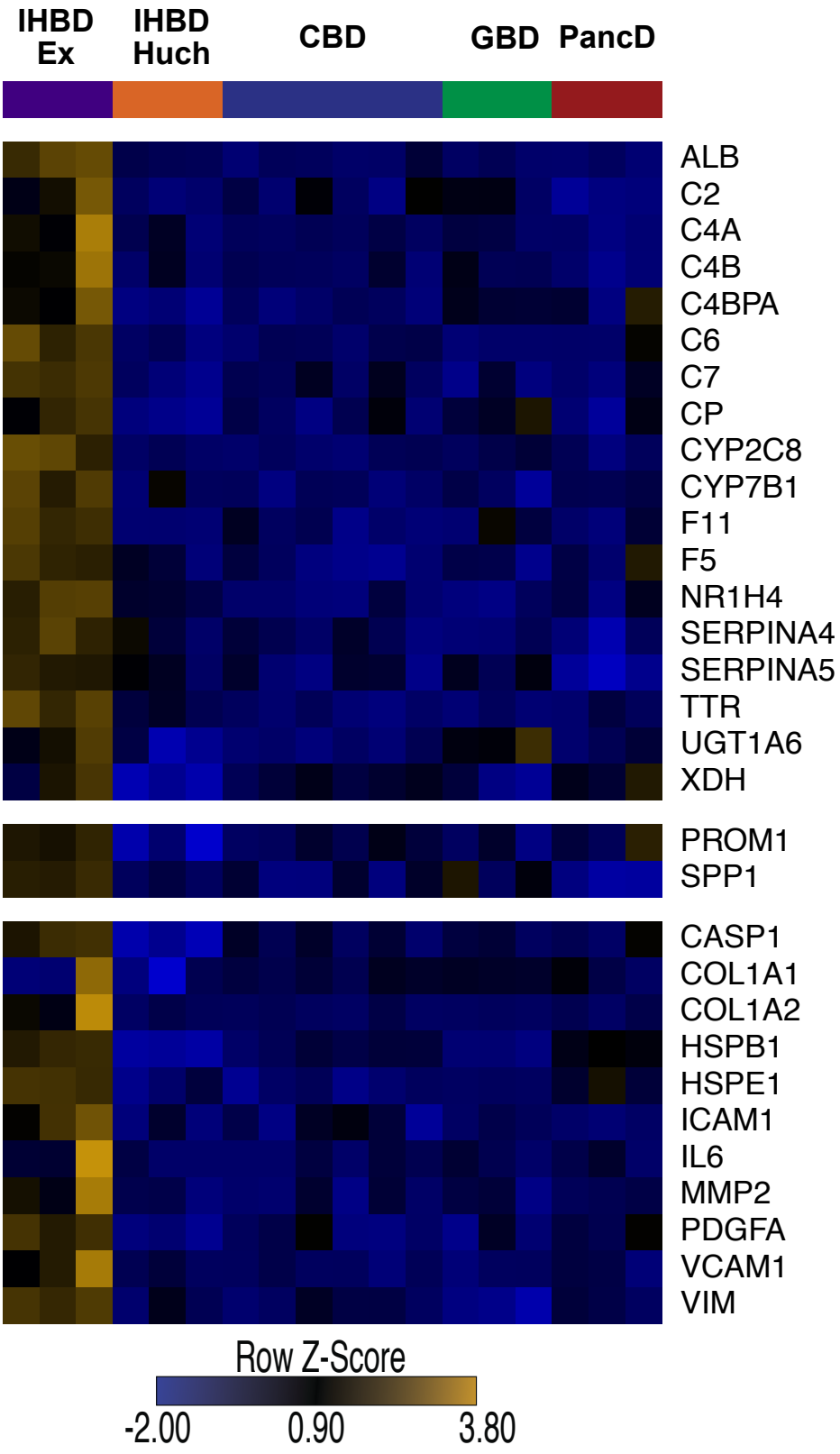


Figure 4.40 Heatmap of selected genes upregulated in IHBD_ex organoids.

Heatmap of selected genes involved in mature hepatocyte function (cluster 1), liver progenitor cells (cluster 2) and cell death/epithelial mesenchymal transition (cluster 3) that were upregulated in IHBD_ex compared to both IHBD_Huch and EHBD organoids.

Next, we sought to more broadly understand the differences that may exist between the IHBD and EHBD organoids. Given the transcriptional signature of the IHBD_ex organoids, and their known proliferative difference, we focused these comparisons on EHBD organoids and IHBD_Huch organoids. 1,393 genes were differentially expressed between EHBD and IHBD_Huch organoids (792 genes upregulated in EHBD and 601 genes upregulated in IHBD_Huch) (**Annex File 11**). Of note, some of these genes also overlapped with genes differentially expressed between EHBD and IHBD_ex organoids (**Figure 4.41**), suggesting EHBD and IHBD organoids are intrinsically different independently of the culture conditions.

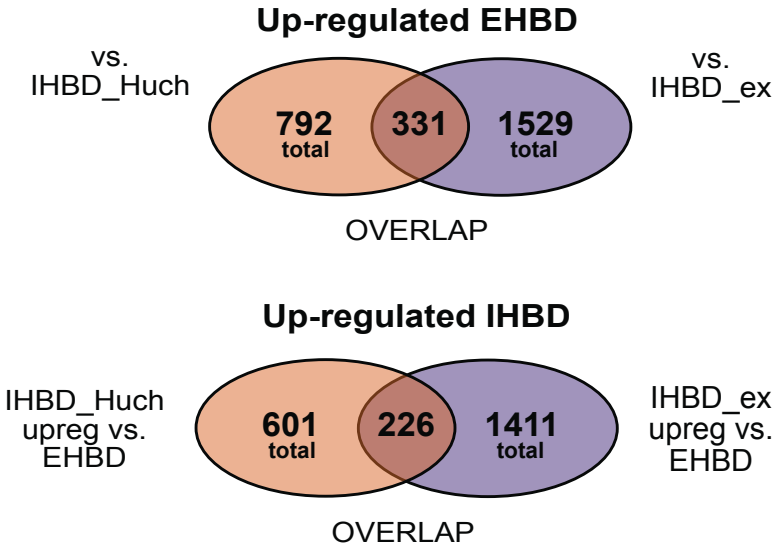


Figure 4.41 Number and overlap of genes differentially expressed between EHBD and IHBD_ex or IHBD_Huch organoids.

Venn diagram showing the number and overlap of genes differentially expressed between EHBD organoids compared to the two IHBD organoid types.

We then examined these genes further and found the 601 genes upregulated in IHBD_Huch organoids were enriched for GO terms involved with developmental processes and cell differentiation (**Figure 4.42, Annex File 12**). Interestingly, these genes included the hedgehog signalling pathway ligands IHH and SHH, which have been shown to promote the maintanence of intrahepatic progenitor cells.[143]

The 792 genes upregulated in EHBD organoids were enriched for GO terms involved in cell cycle processes and DNA replication, including genes such as Cyclin B2/A2/E2 and PCNA. This finding could suggest that proliferation is controlled by different mechanisms between intra- and extrahepatic cholangiocytes and/or that a larger proportion of cells in the EHBD organoids cycle more quickly than IHBD_Huch organoids. Such differences could have broad implications since cholangiopathies are associated with cholangiocyte senescence and thus proliferative capacity could play an essential role in disease development.[131] In sum, these results demonstrate that IHBD and EHBD organoids differ in their proliferative capacity, growth factor requirements and gene expression profile.

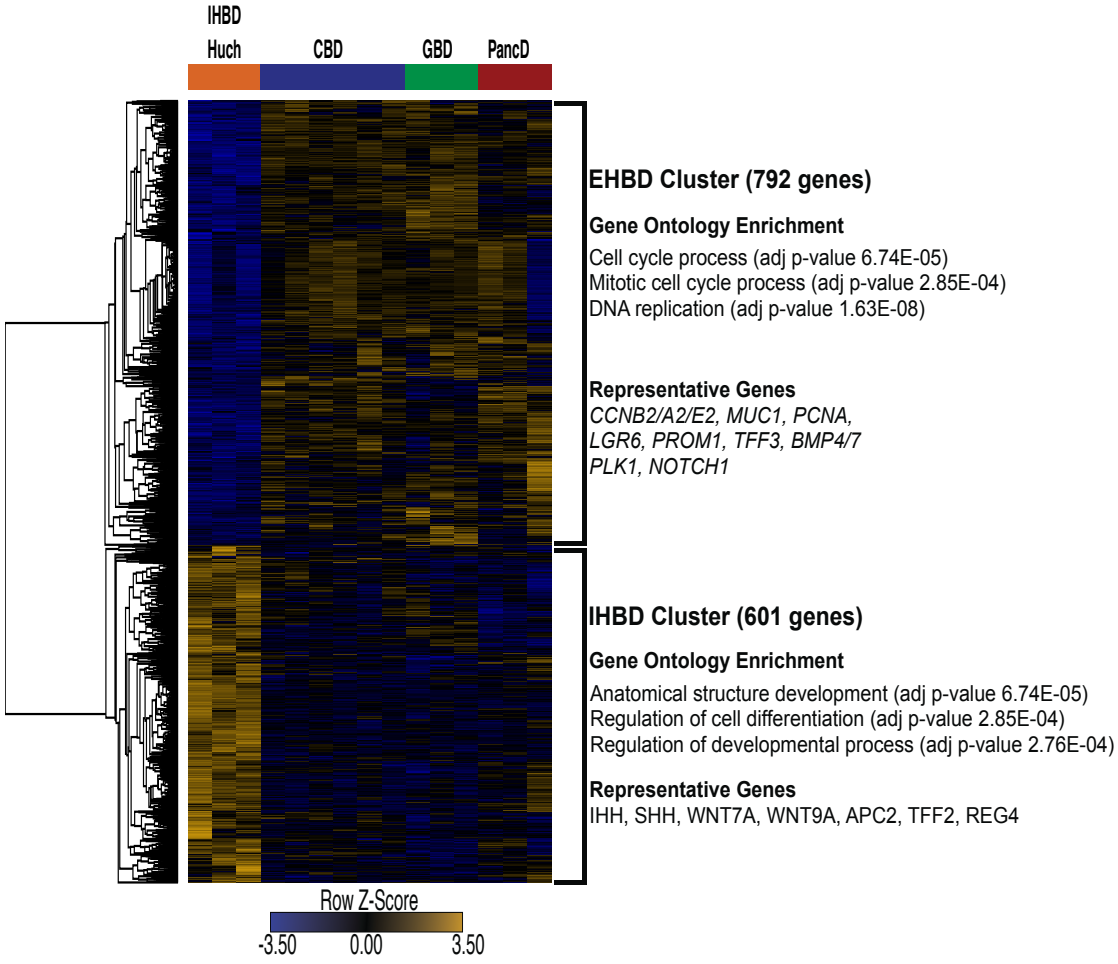


Figure 4.42 Heatmap of the 1,393 genes differentially expressed between EHBD and IHBD_Huch organoids

Heatmap of the 792 genes upregulated in EHBD organoids and the 601 genes upregulated in IHBD_Huch organoids, as well as representative gene ontology terms (biological processes) and examples of genes upregulated in each organoid type.

4.3.8 Exploratory analyses of the transcriptional profile of intrahepatic biliary tissues suggest differences from extrahepatic biliary tissues

As mentioned above, three IHBD primary tissue samples were also sequenced. These samples were obtained during isolation of IHBD organoids. Liver tissue from three donors was frozen in Cell Banker 2 freezing media. The liver tissue was thawed and enzymatically dissociated until only ductal structures remained visible in the digestion media. These ductal structures were manually picked using a p200 pipette, under a microscope, pelleted, and lysed for RNA. Dissociation methods similar to this have been previously employed and shown to enrich for IHBD units with efficiencies as high as 90%.[66] However, in our RNA-Sequencing data, it was noted that the IHBD tissue samples clustered very far from EHBD tissue samples by PCA analysis (**Figure 4.43**).

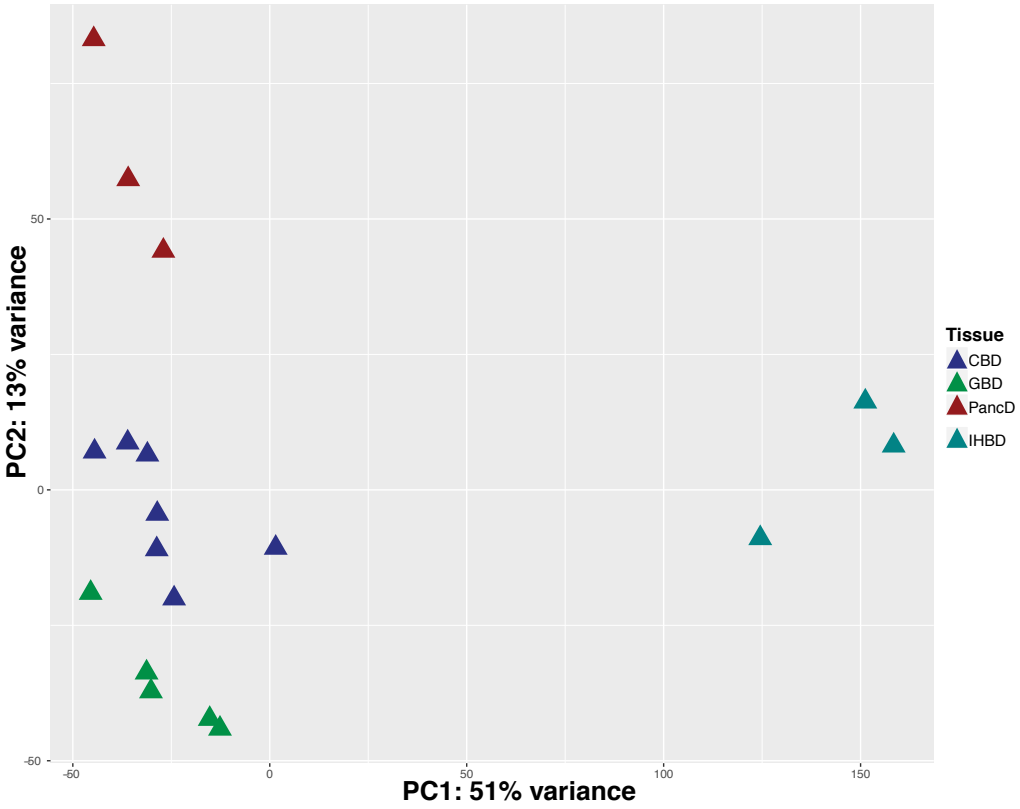


Figure 4.43 Principal component analysis of EHBD and IHBD tissues.

Principal component analysis of variance stabilized counts for the top 5,000 most variable genes between CBD tissue (n=7), GBD tissue (n=5), PancD tissue (n=3) and IHBD tissue (n=3).

Therefore, we further examined the sequencing data for biliary specific markers KRT7, KRT19, SOX9, EPCAM, and HNF1B in the IHBD tissue samples compared to our

EHBD tissue samples. We found that all three of the EHBD tissues (CBD, GBD, and PancD) had significantly higher expression of several of these markers by DGE analysis (KRT19, HNF1B, EPCAM) (**Figure 4.44**). KRT7 and SOX9 were not differentially expressed. We also examined the hepatocyte specific marker ALB and found it was significantly up-regulated in IHBD tissues compared to EHBD tissues. These observations were further confirmed by qPCR analysis and we compared the expression profile of IHBD tissue samples to that of whole liver tissue and to EHBD tissues (**Figure 4.45**).

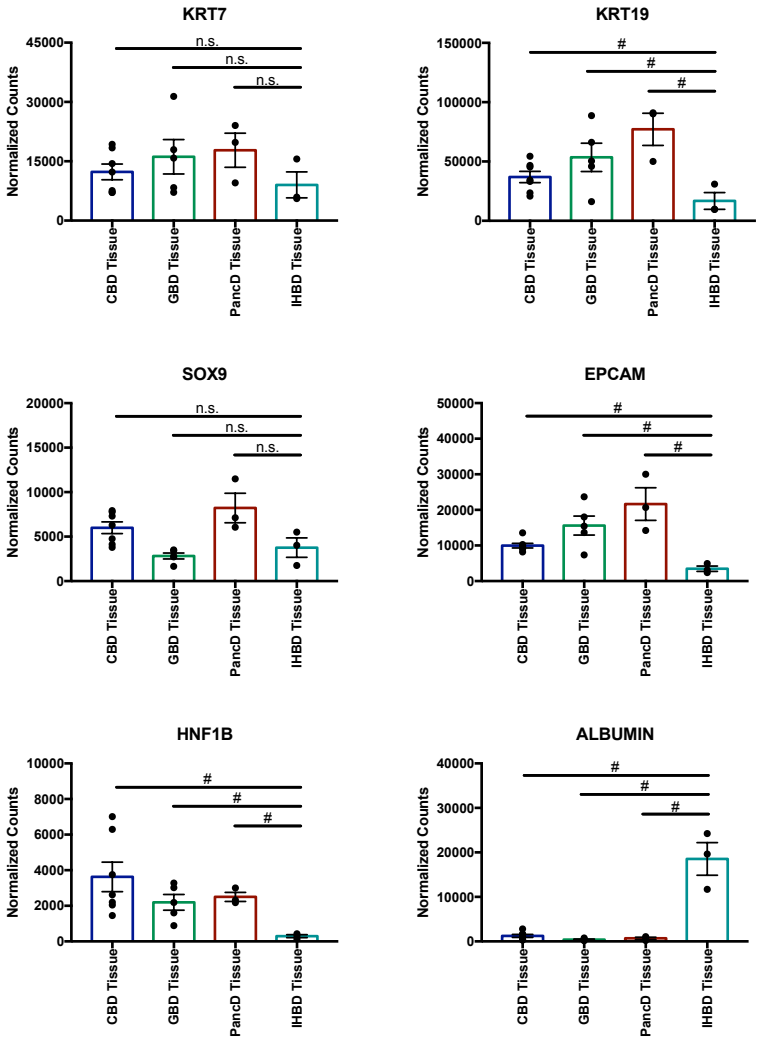


Figure 4.44 RNA-Seq normalized counts for biliary specific markers and liver markers in EHBD and IHBD tissue samples.

DeSeq2 normalized counts for bile duct epithelial specific markers and Albumin. # = false discovery rate (adj p-value) < 0.05 and fold change > 1. n.s.= not significant (adj p-value > 0.05).

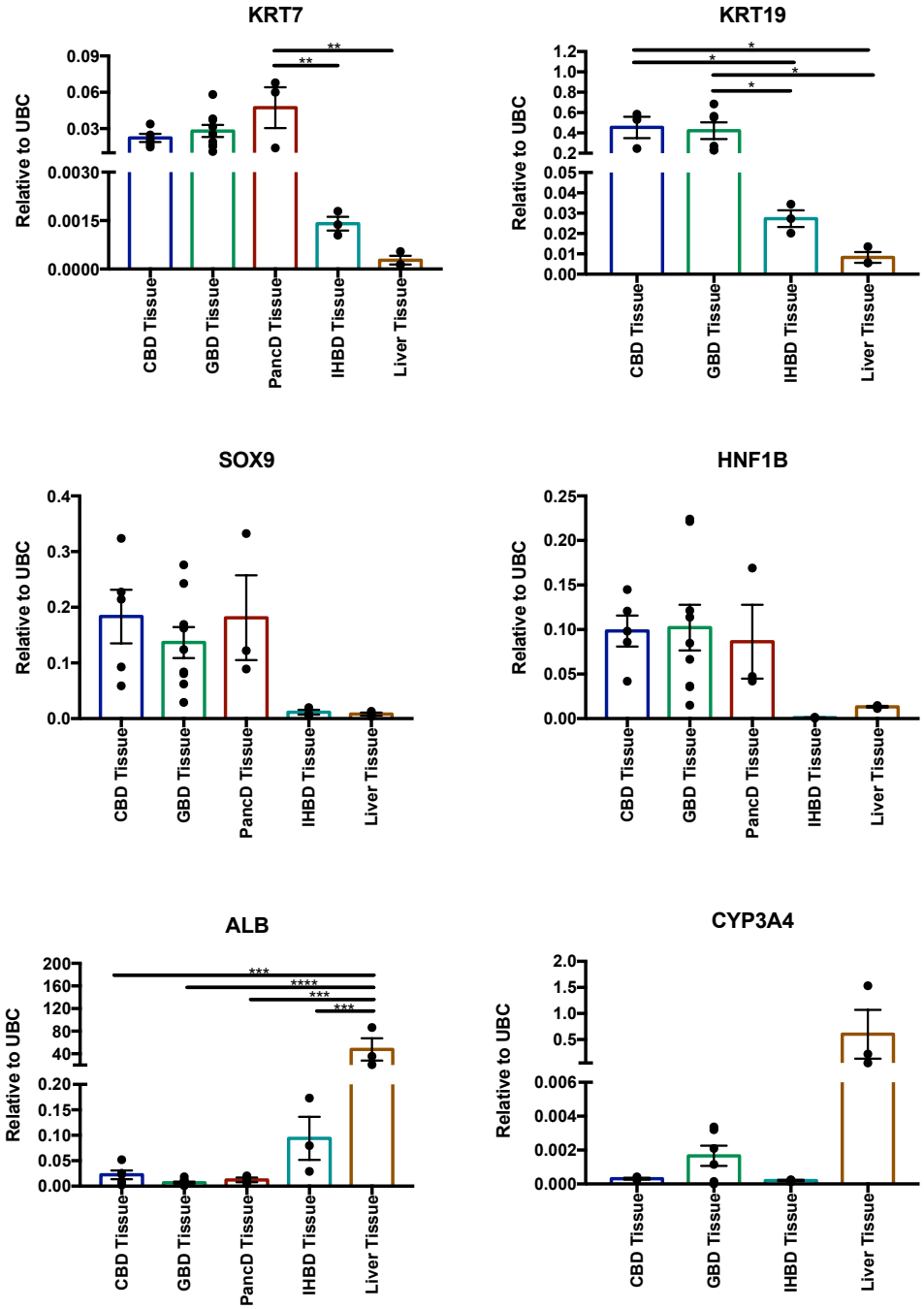


Figure 4.45 Quantitative PCR analysis for biliary and liver markers in EHBD, IHBD, and whole liver tissue samples.

qPCR analysis for biliary markers and hepatocyte markers. Gene expression is normalized to the housekeeping gene UBC and data is plotted as mean and SEM. For KRT7, SOX9, HNF1B, and ALB: CBD tissue (n=5), GBD tissue (n=9), and PancD tissue (n=3). For KRT19 and CYP3A4: PancD tissue (not available), CBD tissue (n=3) and GBD tissue (n=6). All genes: IHBD tissue (n=3) and liver tissue (n=3). *=$p \leq 0.05$, **=$p \leq 0.01$, ***=$p \leq 0.001$, ****$p \leq 0.0001$. If not indicated, comparisons between groups were not significantly different ($p>0.05$)

Taken together, the sequencing and qPCR results suggest enrichment for intrahepatic biliary epithelial cells to some degree, given the higher expression of KRT19 and KRT7, as well as the lower expression of Albumin, in IHBD tissues compared to liver tissue. However, this enrichment was not as high as that achieved for EHBD tissues and IHBD tissues showed lower expression of biliary markers than the EHBD tissue samples. With these limitations in mind, we continued to explore the data to see if anything further could be observed.

As expected, when IHBD and EHBD tissues were compared, a large number of genes were differentially expressed. Between IHBD tissues and GBD tissue, 7,266 genes were differentially expressed (3,930 upregulated in IHBD tissue, 3,336 upregulated in GBD tissue). Between IHBD and CBD tissues 7,000 genes were differentially expressed (3,870 upregulated in IHBD tissue, 3,130 upregulated in CBD tissue). Lastly, between IHBD and PancD tissues 7,915 genes were differentially expressed (4,391 upregulated in IHBD tissue, 3,524 upregulated in PancD tissue). Many of these genes overlapped (**Figure 4.46, Annex File 13**).

GO analysis on overlapping genes upregulated in IHBD tissues compared to the three EHBD tissues showed enrichment for GO terms involved with regulation of developmental process, among others (**Figure 4.47, Annex File 13**). Interestingly, similar GO terms were upregulated in IHBD organoids compared to EHBD organoids in the analyses above. It was also noted that many genes in the NOTCH signalling pathway, a pathway known to be important in biliary fate specification, were upregulated in the IHBD tissues (**Figure 4.48**). It is difficult to conclude whether this expression profile originates from the intrahepatic biliary epithelial cells or from other possible contaminating cell sources. For instance, the NOTCH signalling profile may be originating from periportal mesenchymal cells, as both these and biliary epithelial cells have been shown to express NOTCH pathway genes.[144]

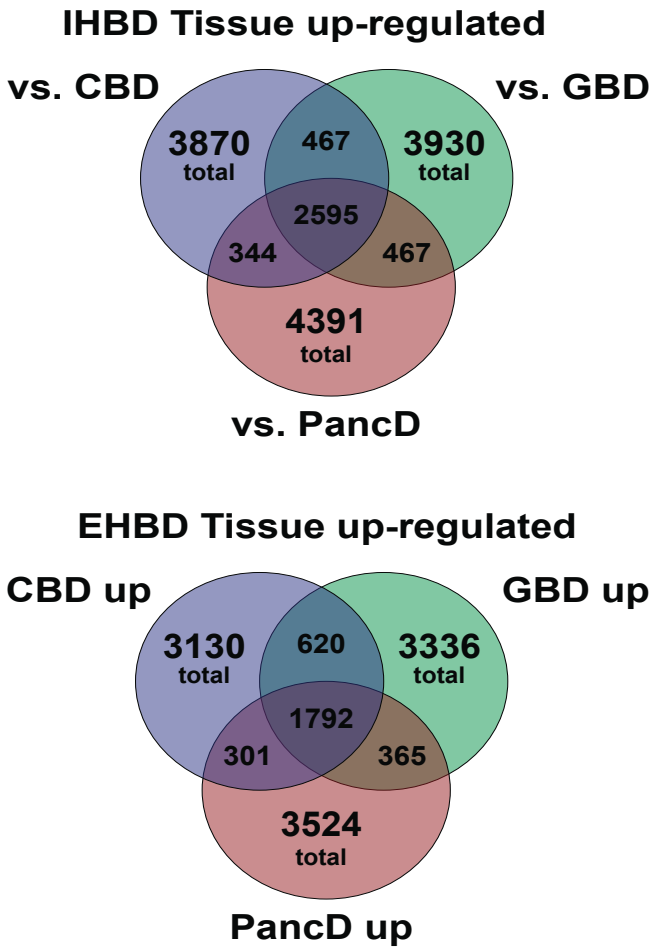


Figure 4.46 Number and overlap of genes differentially expressed between IHBD and EHBD tissues.

Venn diagram showing the number and overlap of genes differentially expressed between IHBD tissue compared EHBD tissues.

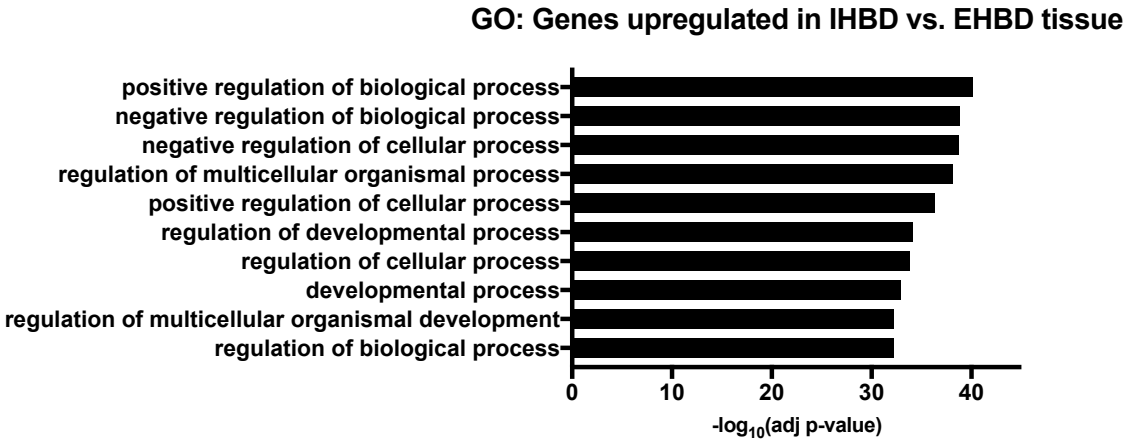
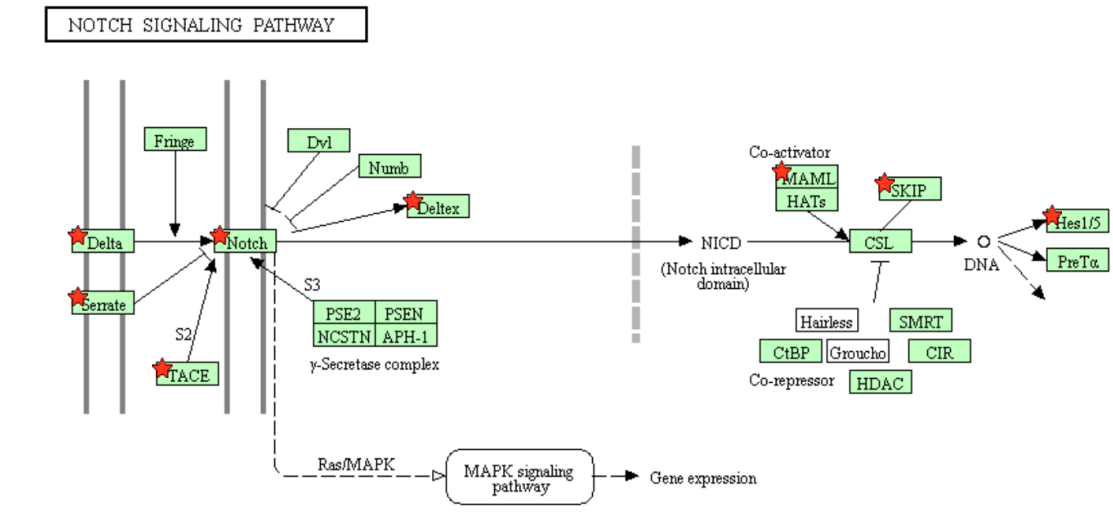


Figure 4.47 Gene Ontology enrichment for genes upregulated in IHBD vs EHBD tissues.

Top ten gene ontology terms (biological processes) which were enriched in the 2,595 genes upregulated in IHBD tissue samples compared to all three of the EHBD tissue regions.



Notch signalling genes: ADAM17, SNW1, DLL1, DLL3, DLL4, DTX1, HES1, HES5, JAG1, JAG2, MAML2, NOTCH3, NOTCH4, HEY1, HEY2

Figure 4.48 NOTCH signalling genes upregulated in IHBD tissues compared to EHBD tissues.

Notch signalling genes upregulated in IHBD tissues compared to all three of the EHBD tissue regions are indicated by red stars on the pathway diagram and listed below the diagram. Pathway diagram generated using DAVID (Database for Annotation, Visualization and Integrated Discovery).[137]

Finally, we examined the expression of the three genes we previously discovered to be tissue specific amongst the three EHBD regions (HOXB2, CDX2, and SOX17) to reveal if any of these genes are also expressed in IHBD tissues or organoids (**Figures 4.49 and 4.50**).

CDX2 remained PancD specific and was significantly higher in both PancD tissue and organoids (**Figure 4.49**). Both IHBD tissue and IHBD_Huch organoids were negative for CDX2 by IF staining (**Figure 4.50**). HOXB2 was not significantly differentially expressed between CBD and IHBD tissue samples in the sequencing data, however, by qPCR it was significantly lower in IHBD tissues. SOX17 was not differentially expressed between GBD and IHBD tissues by both sequencing and qPCR analyses. However, SOX17 was not detectable by IF in IHBD tissue or IHBD_Huch organoids (**Figure 4.50**) suggesting that it is not a marker of IHBDs.

In summary, due to the limitations of the IHBD tissue samples, it is difficult to draw definitive conclusions from the data. However, we were able to show that our previously identified EHBD regional specific genes (CDX2, HOXB2, and SOX17) are likely not markers of IHBD tissues. Further, we demonstrated that manual enrichment of IHBDs is not sufficient to isolate an enriched population of biliary epithelial cells. Lastly, we identified the NOTCH signalling pathway as a potential pathway upregulated in IHBDs, however, further validation of this finding, on a pure population of IHBD biliary epithelial cells, is necessary.

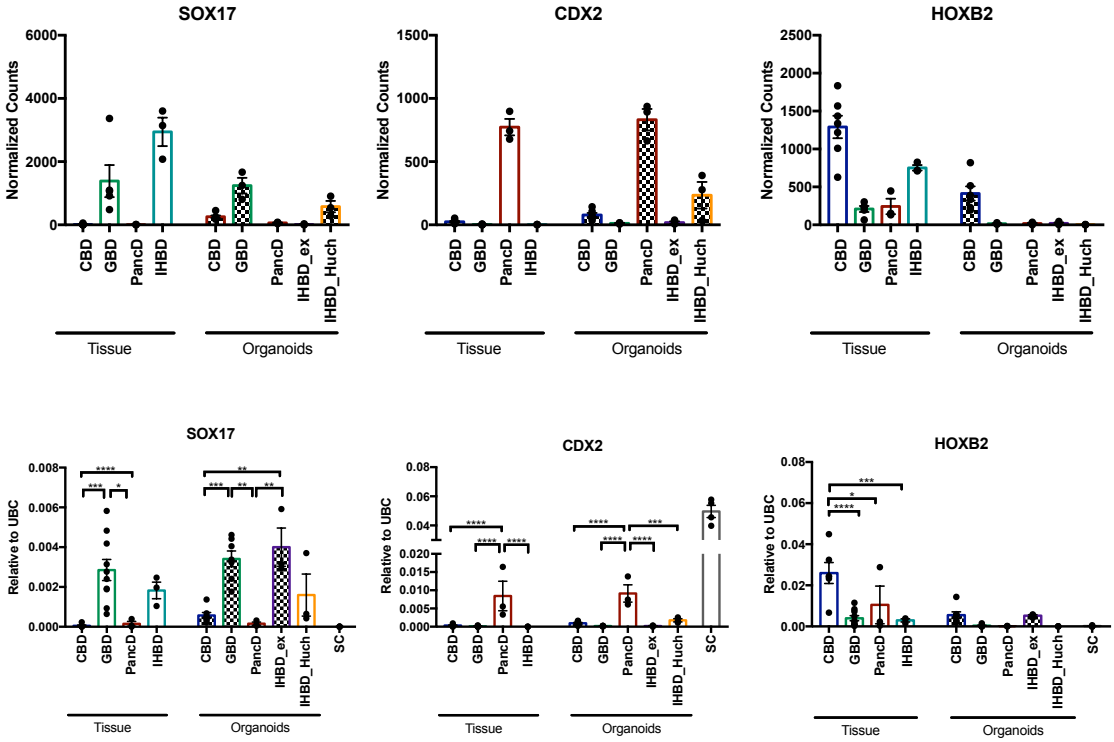


Figure 4.49 RNA-Seq normalized read counts and qPCR validation on IHBD tissues and organoids for three genes previously identified as regional specific markers of EHBD tissues.

RNA-Seq normalized read counts and qPCR validation on IHBD tissue and organoids of three genes previously found to be regional specific in EHBD tissues. For qPCR validation: CBD Tissue (n=6), CBD Organoids (n=7), GBD Tissue (n=10), GBD Organoids (n=7), PancD Tissue (n=3), and PancD organoids (n=3), IHBD tissue (n=3), IHBD_Huch organoids (n=3), and IHBD_ex organoids (n=3). Gene expression is normalized to the housekeeping gene UBC and data is plotted as mean and SEM. Sigmoid colon (SC) organoids (n=4) were used as controls. *=$p \leq 0.05$, **=$p \leq 0.01$, ***=$p \leq 0.001$, ****$p \leq 0.0001$. If not otherwise indicated, comparisons between groups were not significantly different ($p>0.05$)

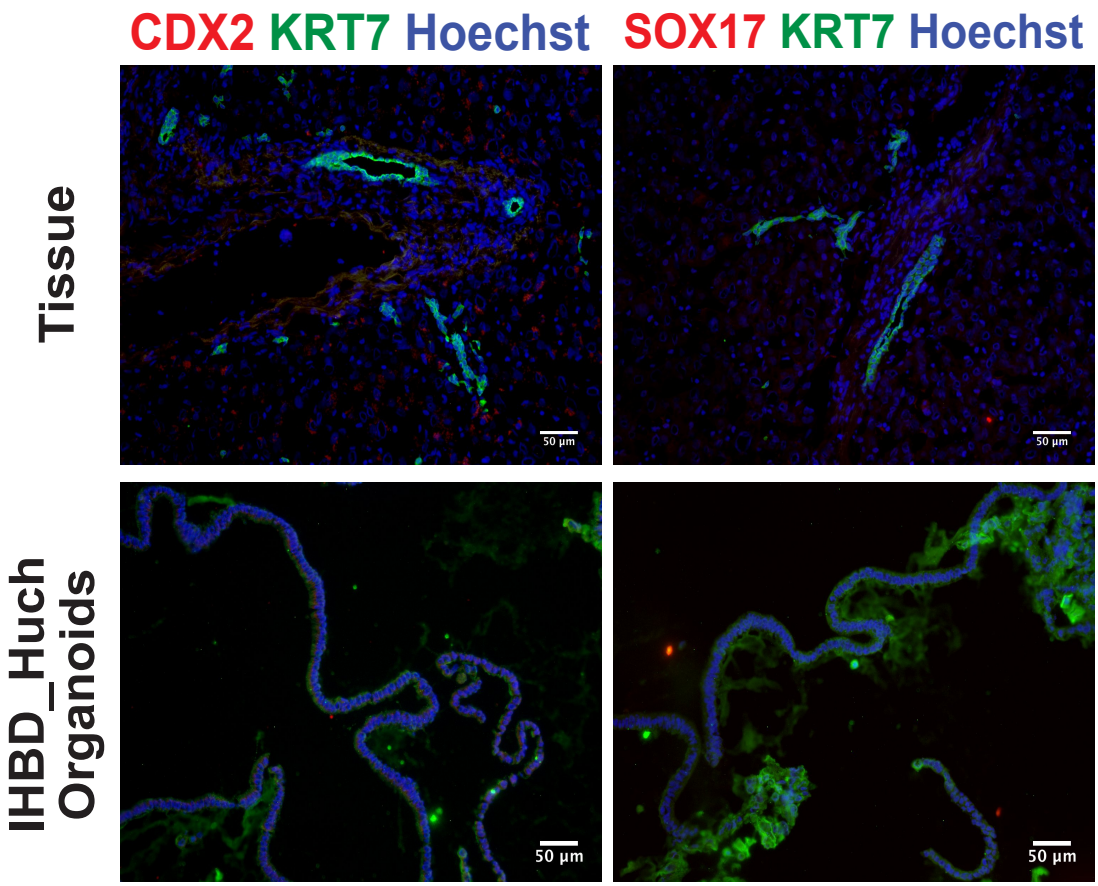


Figure 4.50 Immunofluorescence staining on human liver tissue and IHBD_Huch organoids for CDX2 and SOX17.

Scale bar=50 µm

4.4 Discussion

This chapter presented the results from a comprehensive transcriptional profiling of human biliary tissues and *in vitro* cultured organoids derived from these tissues. Using RNA-sequencing, a transcriptional profile of epithelial enriched samples from three EHBD tissue regions including the CBD, GBD, and PancD was established. These profiles allowed assessment of differences in expression between the three tissues and allowed regional specific markers for these tissues, including HOXB2, CDX2, and SOX17, to be uncovered. Further, the consistency of these results with previously published work was demonstrated. Many of the genes identified as being GBD tissue specific compared to both CBD and PancD tissues in our own dataset, were also previously reported to be GBD tissue enriched, even when compared to 27 other unrelated human tissues. Our findings on SOX17 and CDX2 have precedence in the literature. In the case of SOX17, some reports in the mouse suggest it to be a marker of only GBD tissue, however, others suggest it is also expressed in bile duct tissue.[53,94] To our knowledge, this finding has not been confirmed before in human tissues and we show that SOX17 is only expressed in GBD tissue and not expressed in CBD, PancD, or IHBD tissue at the protein level. Two reports of CDX2 expression in PancD tissues exist, however, it was not known until now that it is a marker of only PancD tissue and not other regions of the EHBD biliary tree.[145,146] It was also observed that even genes considered to be biliary specific genes, such as ABCC2, SLC10A2 (ASBT), and SCTR, were differentially expressed between the three EHBD regions, with GBD tissue having upregulated expression of ABCC2 and SLC10A2, while SCTR was upregulated in PancD tissue.

Additionally, we were able to gain insight into functional differences these tissues may have. Most notably, we discovered that GBD tissue has higher expression of genes associated with active metabolic processes involved in bile modification, including xenobiotic metabolism, lipid metabolism, and cholesterol metabolism, when compared to both CBD and PancD tissues. Ultimately, it appears from the data, that each of these three tissue regions display a transcriptional profile in line with their function in bile transport, while at the same time maintaining a key underlying expression of essential cholangiocyte-specific genes (i.e. CFTR, HNF1B, KRT7, KRT19). The GBD acts as the primary site of bile storage, and as such, displays a transcriptional profile consistent with having the predominant role in bile modification. While the CBD acts as a conduit for

bile and is in contact with bile for a shorter period of time and after it leaves the GBD. This may explain why the CBD tissue had lower expression levels of key bile modification pathways when compared to GBD tissue, as when the bile reaches the CBD it has already been modified extensively in the GBD. Additionally, the PancD, which is never in contact with bile, has downregulated and/or no expression of these bile modification pathway genes compared to both the CBD and GBD tissues. Instead, the PancD tissue appears to be enriched for genes involved in protein synthesis and RNA processing. This enrichment however, could be explained by contaminating acinar cells, as they produce a large amount of secreted proteins and require a high level of translational machinery. It is intriguing to consider whether the epithelial cells themselves impart this functional difference between tissue regions or if the anatomical environment itself, such as the higher bile concentration in the GBD, imposes these functional differences onto the epithelial cells.

Also, interesting to consider, are the effects of GBD removal, or cholecystectomy. Cholecystectomy is one of the most common surgeries performed worldwide and the GBD is considered as a non-essential organ as patients usually tolerate cholecystectomy with limited to no long-term side effects.[51] Following cholecystectomy, however, it has been shown that many biological alterations do occur. These include things such as a decrease in circulating levels of FGF19, increased enterohepatic recycling of bile acids, alterations to the composition of the bile acid pool, and even possible changes in the gut microbiota. There have even been reports on the association of cholecystectomy with an increased risk of developing metabolic syndrome (type 2 diabetes, high cholesterol/triglycerides) and non-alcoholic fatty liver disease. This risk was found to be independent of gallstone disease.[51] These associations are very interesting, as our own data suggests that GBD epithelial cells likely play an active role in the metabolic processes that are perturbed in these disease states. Therefore, cholecystectomy may in fact be removing a reservoir of cells that participate in crucial metabolic functions. Interestingly, our data also suggested that, when compared to the PancD, CBD tissue may serve as a site for some of these metabolic functions, although to a lesser degree than GBD tissue. It would be interesting to assess in future studies how CBD epithelial cells adapt to cholecystectomy and whether the CBD or other regions of the biliary tree are capable of taking on some of the functional roles that the GBD epithelium appears to be important in maintaining.

One key limitation to this RNA sequencing data, which must be mentioned, is that the EHBD tissue samples were derived from epithelial enriched samples obtained by mechanical dissociation and were not purified beyond this. This method, as we demonstrated in Chapter 1, does enrich for biliary epithelial cells with only 5-10% of the cells not expressing KRT7/KRT19. However, in the RNA-sequencing data, we could observe expression of genes not typically associated with biliary epithelial cells. For example, we observed expression of blood cell markers such as CD4 and Haemoglobin in the three EHBD tissue samples. This limitation was most apparent in the PancD tissue samples and we found the PancD tissue samples had significantly higher expression of several acinar and endocrine-specific genes. This suggests that the PancD tissue samples may have contained a low-level of contaminating pancreatic parenchymal tissue. This does lead to limitations in the interpretations that can be drawn from the genes upregulated in PancD tissue samples and this data should be considered cautiously as they may not originate from the pancreatic ductal cells. However, even given this limitation, we were still able to uncover PancD specific gene signatures, including CDX2, and gain insight into genes that are upregulated in GBD/CBD tissue in comparison to PancD tissues.

In this chapter, we also characterized the transcriptomic profile of organoids derived from EHBD and IHBD tissue sources. It was found that EHBD organoids are different from their tissue of origin but that a small number of regional tissue specific markers are maintained in the organoids. We also discovered that the organoids upregulate expression of cell cycle genes, WNT signalling genes, and genes that have previously been associated with intestinal adSCs. The organoids also down-regulate many functional markers of mature biliary cells. These results continue to support the idea that the EHBD organoids likely contain a population of cells with a stem/progenitor-like phenotype, consistent with our earlier characterizations of the organoids in Chapter 3. Further, we demonstrated that IHBD_Huch and IHBD_ex organoids have unique transcriptional profiles from that of the EHBD organoids and also from each other. However, the biological significance of these differences was difficult to interpret given the proliferative differences observed in the IHBD_ex organoids and the different media components needed to culture IHBD_Huch organoids. This limitation causes difficulties in interpreting weather differences in the organoid transcriptional profiles are due to IHBD versus EHBD differences, or if they are due simply to differences in proliferation rate and/or media composition. Despite this limitation, we were able to discover that our extrahepatic

organoid culture conditions may be inducing IHBD_ex organoids to differentiate towards a hepatocyte-fate or undergo EMT. Interestingly it has been shown that WNT signals from macrophages in the liver act to push intrahepatic liver progenitor cells towards a hepatocyte fate during tissue injury.[147] It is possible this may explain the differences we observed in IHBD organoids cultured in media containing CHIR 99021.

Lastly, we considered sequencing data from IHBD tissue samples. However, it was found that these samples were likely contaminated by other cell types including hepatocytes and possibly mesenchymal cells from the liver, making it difficult to draw definitive conclusions from the data. However, despite these limitations in the samples, we were still able to examine the expression of CDX2, SOX17, and HOXB2 in IHBD tissues and organoids. We found these genes are likely not expressed to a significant degree in the IHBD compartment and that they are truly unique tissue-specific markers for different regions of the extrahepatic biliary tree.

In summary, we feel these results are significant and have provided the first evidence that the transcriptional profile of biliary tissues differs throughout the human extrahepatic biliary tree. Furthermore, we found that unlike other organoid systems, EHBD organoids are unique from their tissue of origin. Lastly, we show that IHBD and EHBD organoids are not equivalent and that differences exist in their transcriptional profiles.

5 ASSESSING THE DIFFERENTIATION CAPACITY OF HUMAN BILIARY ORGANOIDS

5.1 Statement of Source

The results presented here are largely based on experiments presented in the following first author manuscript written by the author of this dissertation. Therefore, some parts have been taken *verbatim* or with only minor changes from this source.

Rimland CA, Tilson S, Morell C, Tomaz R, Lu WY, Adams S, Georgakopoulos N, Otaizo-Carrasquero F, Myers T, Sun HW, Gieseck RL, Sampaziotis F, Tysoe O, Wesley B, Oniscu GC, Hannan NRF, Forbes S, Saeb-Parsy K, Wynn TA, Vallier L. (2018) Regional differences in human biliary tissues and corresponding *in vitro* derived organoids. Manuscript in preparation.

The experiments presented in this chapter were performed in collaboration with several individuals and this is noted in the figure legends.

In particular, Ms. Samantha Tilson assisted in performing the differentiation experiments on IHBD_Huch and IHBD_ex organoids. While Ms. Simone Adams assisted in the screening experiments on EHBD organoids while under the supervision of the author of

this dissertation. Ms. Olivia Tysoe and Mr. Kourosh Saeb-Parsy performed the liver capsule injections. All data analysis, figure preparation and interpretation of these experiments was performed by the author of this dissertation.

5.2 Introduction

As introduced in Chapter 1.4, there are mixed reports in the literature on the differentiation capacity of biliary organoids. It has been shown that intrahepatic biliary cells *in vivo* and intrahepatic bile duct derived organoids likely contain a bipotent progenitor population with the capacity to express hepatocyte-specific markers under certain conditions. However, the ability of extrahepatic biliary cells and organoids to differentiate is still unclear. Several reports suggest it may be possible, while others have shown extrahepatic cells are limited in their differentiation capacity.[111,113,116,122] Therefore, in this chapter we sought to assess the differentiation capacity of our EHBD organoid cell lines. As we presented in the previous two chapters, these EHBD organoids share many characteristics of adSCs with IHBD organoids such as the expression of key markers LGR5 and PROM1. However, as our analyses in Chapter 3 and 4 demonstrated, the EHBD organoids are distinct from IHBD organoids *in vitro* suggesting that their differentiation capacity may also be divergent.

5.3 Results

5.3.1 Extrahepatic biliary organoids do not demonstrate differentiation capacity towards a hepatocyte fate when using a previously published differentiation protocol

In the course of our experiments detailed in the previous two chapters, we were able to assess EHBD organoids for several characteristics of stem/progenitor cells. EHBD organoids demonstrated long-term proliferative ability and expressed markers traditionally associated with adult stem/progenitor cells. However, stem/progenitor cells should have the ability to differentiate. To assess this, we first took advantage of the differentiation protocol previously published by Huch et al (2015) to assess the capacity of EHBD organoids to express hepatocyte-specific markers, potentially indicating a capacity to differentiate (**Figure 5.1**). IHBD_Huch organoids from three donors and IHBD_ex organoids from one donor were used as positive controls.

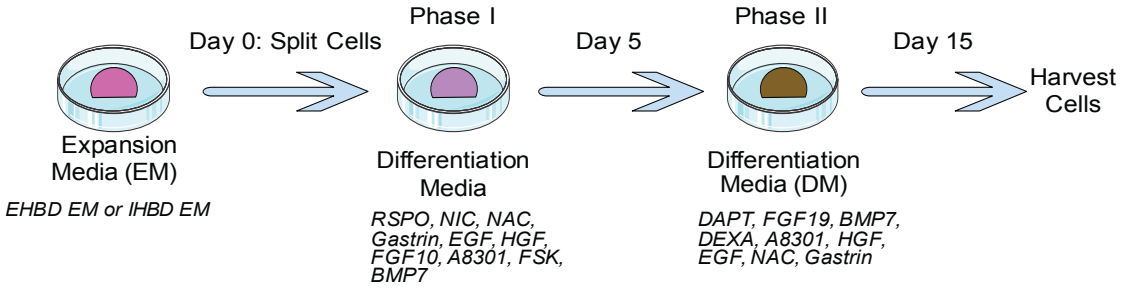


Figure 5.1 Diagram depicting the differentiation protocol described by Huch et al (2015) used to assess the differentiation capacity of EHBD and IHBD organoids.

As previously reported, IHBD_Huch organoids, under differentiation conditions, showed significantly reduced expression of the stem cell markers LGR5 and PROM1 by qPCR (**Figure 5.2**). KRT19 continued to be expressed while SOX9 expression decreased. Induction of a low-level expression of hepatocyte-specific markers such as TTR, CYP3A4, and ALB was also observed in IHBD_Huch organoids by qPCR. However, as previously reported, their expression was significantly lower than expression levels in primary human hepatocytes. IHBD_ex organoids demonstrated identical trends (**Figure 5.3**) in expression indicating the conditions in which IHBD organoids were isolated and maintained did not impact their ability to express these hepatocyte-specific markers.

Next, the effect of this differentiation protocol on EHBD organoids from all three tissue regions was assessed. When subjected to the differentiation protocol, EHBD organoids, similarly to IHBD organoids, had significantly decreased expression of LGR5, PROM1, and SOX9, with stable expression of KRT19 (**Figure 5.2**). However, EHBD organoids showed no induction of the hepatocyte-specific markers TTR, ALB, or CYP3A4. These results were confirmed at the protein level showing ALB expression exclusively in IHBD organoids (**Figure 5.4 and Figure 5.5**).

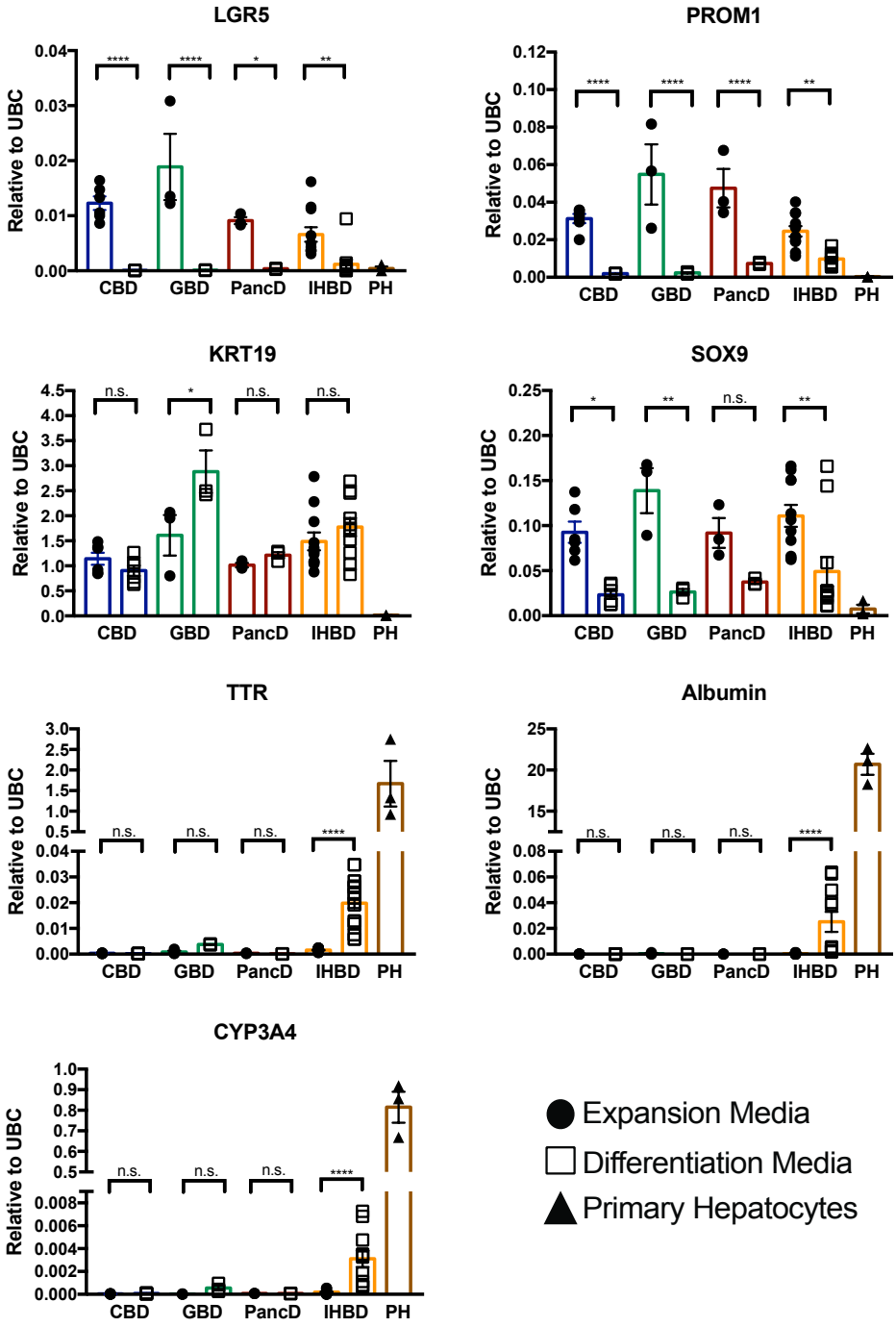


Figure 5.2 Extrahepatic bile duct organoids do not display differentiation capacity towards a hepatocyte fate when compared to intrahepatic bile duct organoids.

qPCR analyses showing expression of biliary and hepatocyte markers in IHBD_Huch, CBD, GBD, or PancD organoids grown in expansion media (EM) or differentiation media (DM) (n=3-11 independent experiments, n=1-3 donors per tissue region). IHBD samples were generated in collaboration with Ms. Samantha Tilson. Primary hepatocytes (PH, n=3) were used as controls. Gene expression is normalized to the housekeeping gene UBC and data are plotted as mean and SEM. *=$p \leq 0.05$, **=$p \leq 0.01$, ***=$p \leq 0.001$, ****$p \leq 0.0001$.

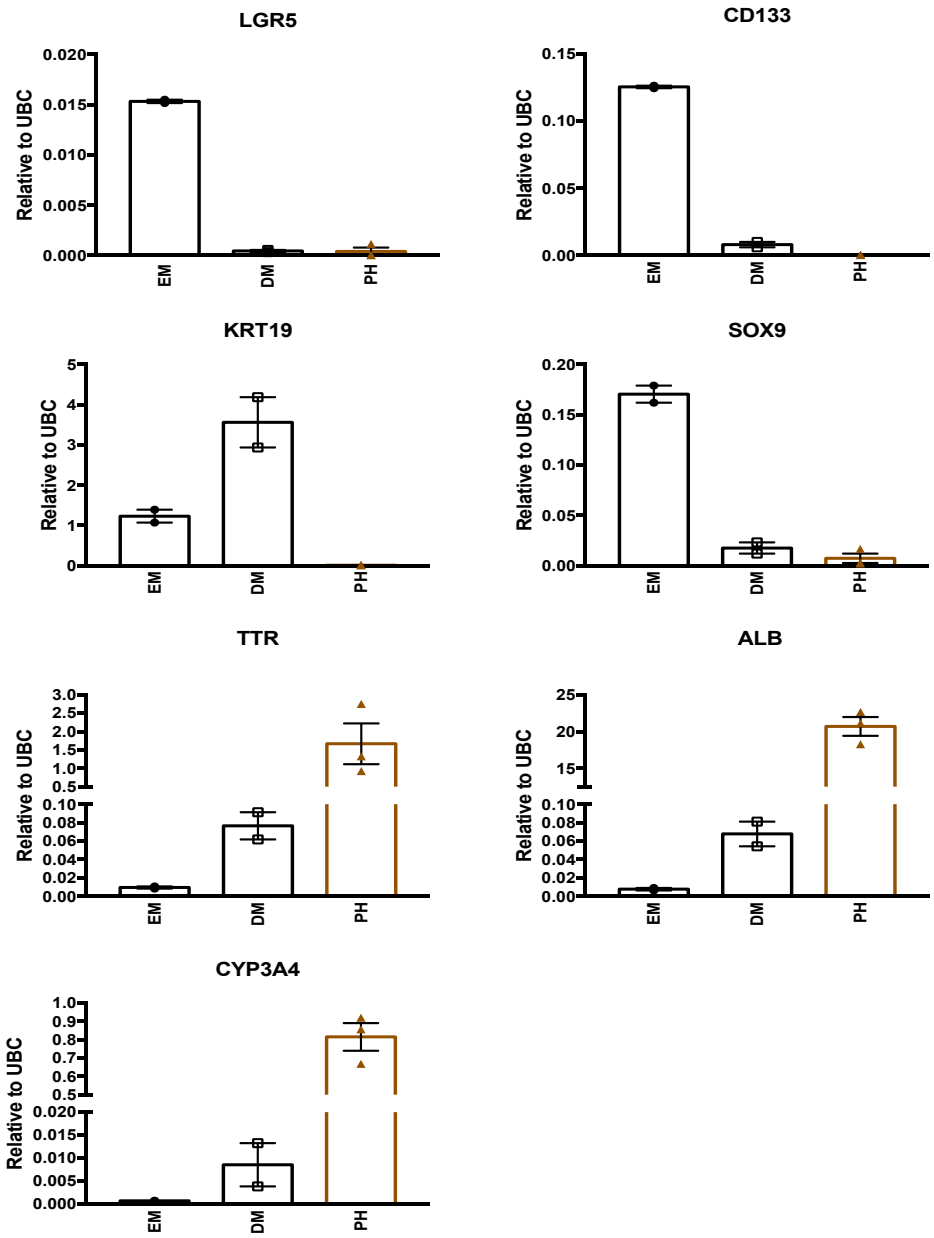


Figure 5.3 IHBD organoids grown in extrahepatic culture conditions up-regulate hepatocyte specific markers under differentiation conditions.

qPCR analysis for stem cell, biliary, and liver genes in IHBD_ex organoids (n=2 independent experiments, n=1 donor line) in either expansion media (EM) or differentiation media (DM). IHBD samples were generated in collaboration with Ms. Samantha Tilson. Primary human hepatocytes (n=3) were used as positive controls. Gene expression is normalized to the housekeeping gene UBC and data is plotted as mean and SEM.

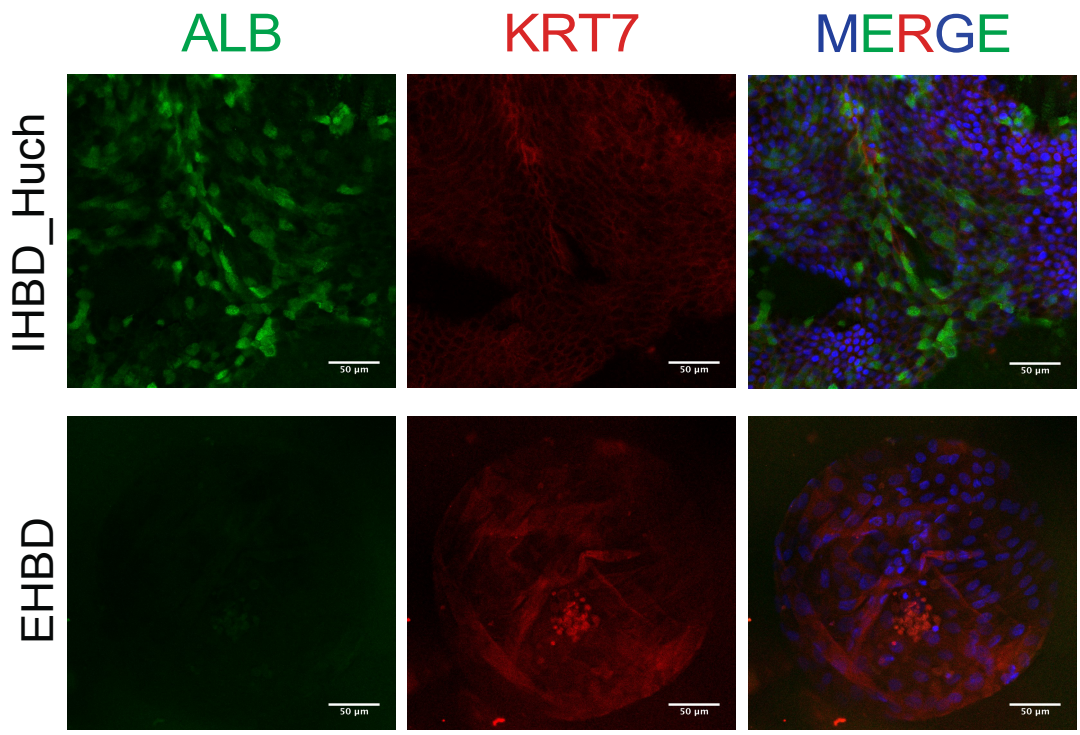


Figure 5.4 Representative immunofluorescence images for Albumin and KRT7 on IHBD_Huch and EHBD organoids cultured in differentiation conditions.

IF images for Albumin and KRT7 of either IHBD_Huch or EHBD organoids cultured for 15 days in the differentiation conditions described by Huch et al (2015).

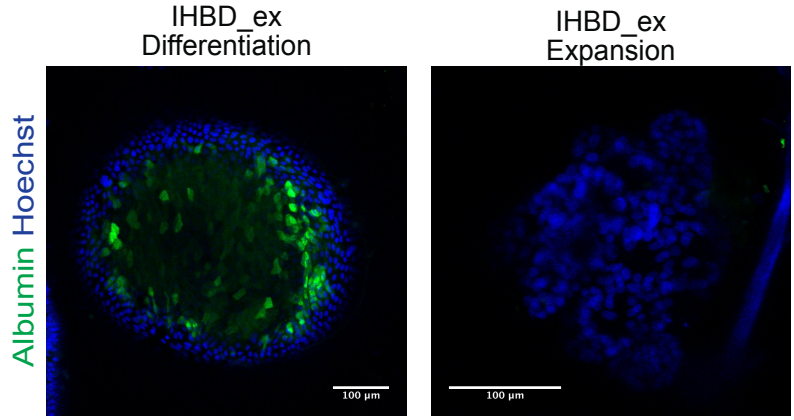


Figure 5.5 Immunofluorescence for Albumin on IHBD_ex organoids cultured in differentiation conditions or expansion conditions.

IF images of IHBD_ex organoids cultured for 15 days in the differentiation conditions described by Huch et al (2015) or alternatively in EHBD organoid expansion media. Experiment performed by Ms. Samantha Tilson.

5.3.2 Screening various culture conditions on extrahepatic bile duct organoids to assess differentiation capacity.

Given the negative results for EHBD organoids, we decided to screen different culture conditions capable of inducing expression of hepatocyte-specific markers. A wide variety of growth factors and small molecule inhibitors aimed at perturbing signalling pathways important in hepatocyte development and liver regeneration were tested (**Figure 5.6, Figure 5.7, and Table 5.1)**. None of these compounds appeared to significantly increase the expression levels of hepatocyte-specific markers in EHBD organoids. qPCR results for stem cell, biliary, and hepatocyte specific genes for single factors screened on at least two independent donor cell lines are shown in **Figure 5.7**.

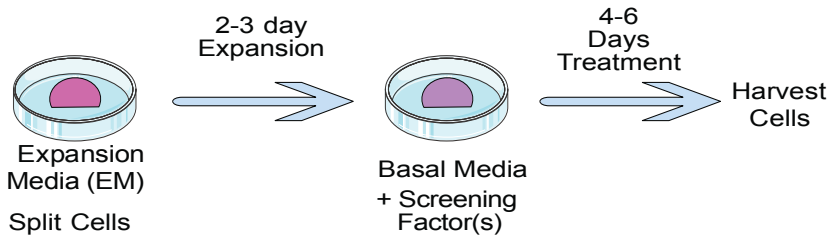


Figure 5.6 Diagram depicting the methods employed for screening experiments performed on EHBD organoids.

EHBD organoids were split and cultured in EHBD expansion media for 2-3 days. The EHBD organoids were then switched into basal media only (ADF+) plus one or a combination of screening factors. The cells were treated for 4-6 days and then harvested for RNA.

Additionally, even when combinations of these factors were tested together, a consistent increase in the hepatocyte markers ALB, CYP3A4, or TTR was not observed (**Figure 5.8**). Interestingly, some of the combinations (i.e. CHIR/DAPT, CHIR/DAPT/SB, CHIR/DAPT/SB/BMP4, and CHIR/DBZ/SB/BMP4) did increase the level of Albumin and CYP3A4 significantly in the EHBD cells compared to the expansion media. Despite being statistically significant, this was overall a low-level induction and in the case of Albumin, induction was magnitudes lower than what was previously observed for IHBD organoids cultured in the differentiation conditions described by Huch et al (2015).

Compound	Pathway	Proliferation Impact	Hepatocyte Markers
Activin (50ng/mL)	TGFβ Activation	-	n.s.
SB 431542 (10µM)	TGFβ inhibition	-	n.s.
A 83-01 (5µM)	TGFβ inhibition	+	n.s.
BMP4 (25ng/mL)	BMP activation	-	n.s.
BMP7 (25ng/mL)	BMP activation	-	n.s.
Noggin (200ng/mL)	BMP inhibition	+	n.s.
Forskolin (10µM)	cAMP activation	+	n.s.
CHIR (3µM)	Wnt activation	- (alone) + (with RSPO)	n.s.
Wnt3a Cond. Media (30%)	Wnt activation	+/o	n.s.
DKK (100ug/ml)	Wnt inhibition	-	n.s.
DAPT (10µM)	Notch Inhibition	-	n.s.
DBZ (10µM)	Notch Inhibition	-	n.s.
HGF (50-100ng/mL)	HGF signaling	+ (caused 2D growth in matrigel)	n.s.
Oncostatin-M (100ng/mL)	OSM signaling	+	n.s.
FGF10 (100ng/mL)	FGF signaling	+	n.s.
FGF7 (100ng/mL)	FGF signaling	+	n.s.
FGF2 (24ng/mL)	FGF signaling	-	n.s.
FGF19 (100ng/mL)	FGF signaling	o	n.s.
Dexamethasone (30 µM)	Glucocorticoid signaling	-	n.s.
Hepatozyme Media	n/a	-	n.s.

+ = increased, o = neither increased or decreased, - = decreased, n.s. = not significant

Table 5.1 List of screening compounds tested to assess differentiation capacity of EHBD organoids.

In the case of CYP3A4, this induction was not seen alongside any significant increases in Albumin or TTR expression. It is possible that this up-regulation of CYP3A4 could be explained by cellular stress as the EHBD organoids did not survive well in these conditions and CYP3A4 expression has been reported to increase in cells undergoing oxidative stress.[148]

Ultimately these results, and those described in the previous section, suggest that only biliary cells located within the liver have the capacity to express hepatocyte markers. This suggests a fundamental divergence in terms of the capacity of differentiation between different regions of the biliary tree.

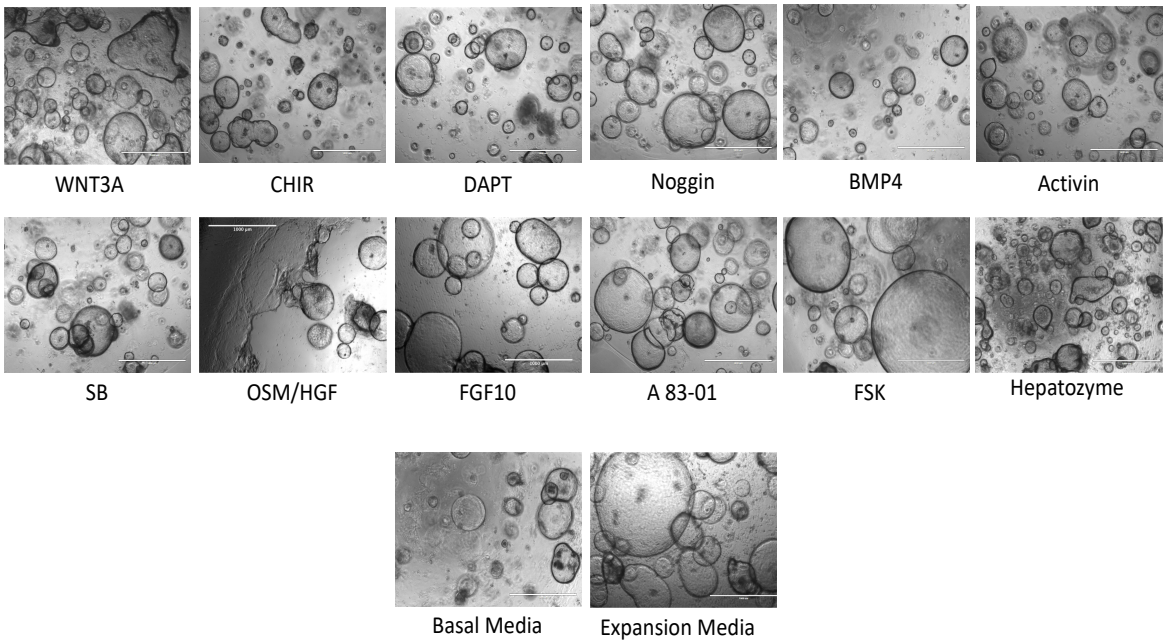


Figure 5.7 Images of organoids treated with a single screening factor for 4-6 days.

Scale bar= 1000 um

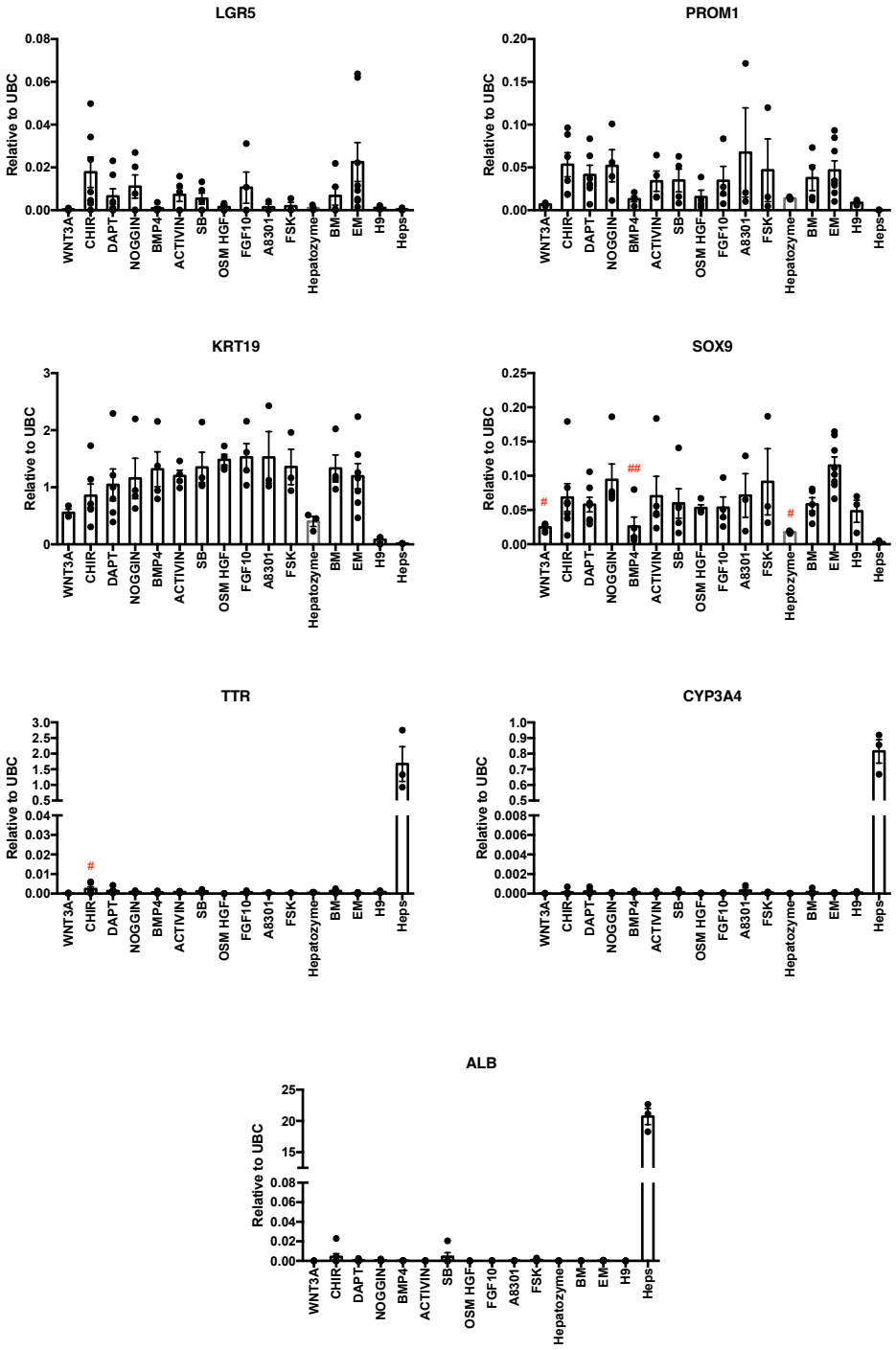


Figure 5.8 Screening of single factors on EHBD organoids to assess for differentiation capacity

qPCR analyses of extrahepatic bile duct organoids (n = 3-7 donor cell lines per condition, except for TTR where A8301 and FSK conditions only have n=2) treated with a single screening factor for 4-6 days. H9 embryonic stem cells (H9, n=3) and primary hepatocytes (Heps, n=3) were used as controls. Gene expression is normalized to the housekeeping gene, UBC, and data is plotted as mean and SEM. EM = extrahepatic bile duct expansion media. BM = basal media alone. # = $p \leq 0.05$, ## = $p \leq 0.01$ for the screening condition compared to EM.

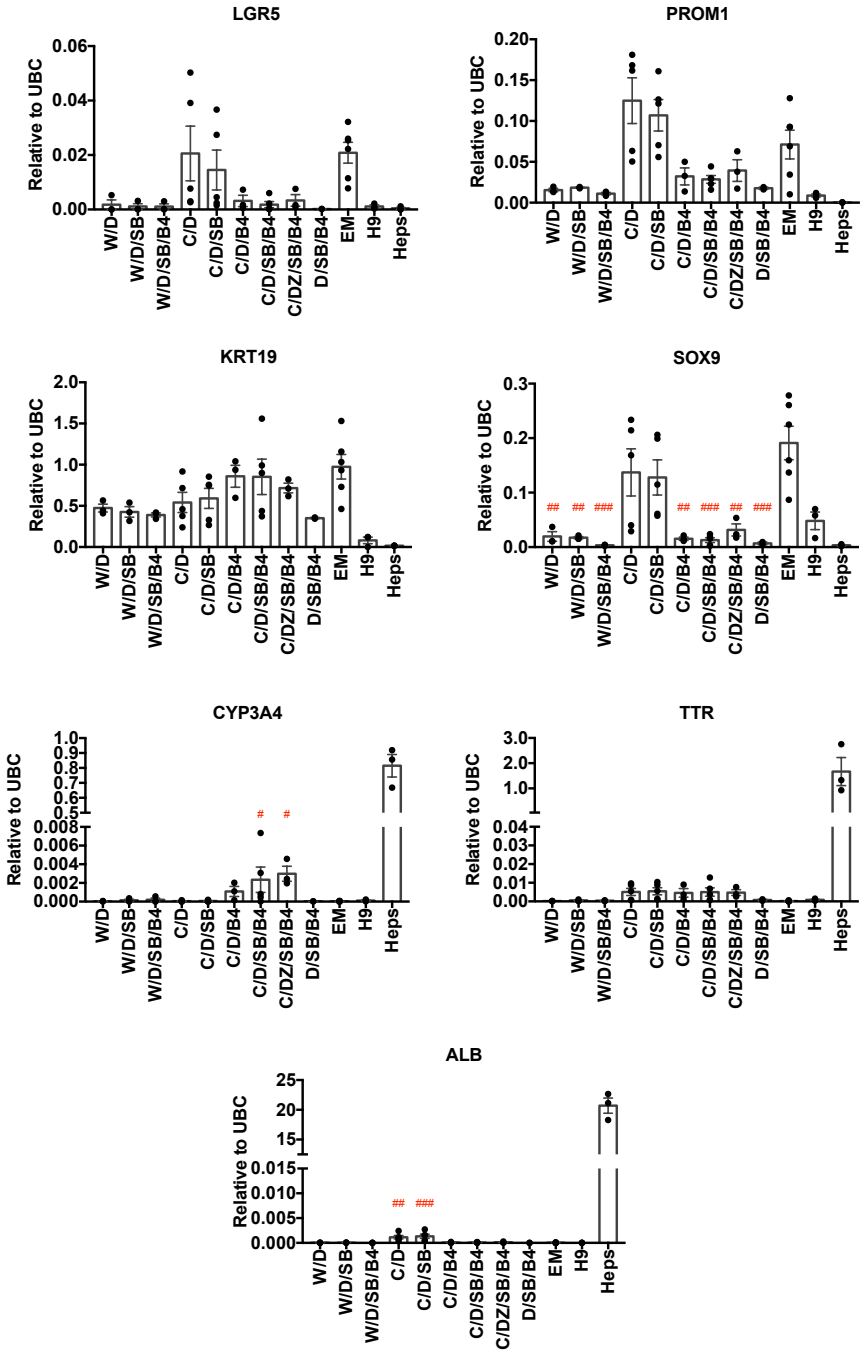


Figure 5.9 Screening combinations of factors on EHBD organoids to assess for differentiation capacity

qPCR analysis of extrahepatic bile duct organoids (n = 3-7 donor cell lines per condition, except for the D/SB/B4 condition which only has n=2) treated with a combinations of screening factors for 4-6 days. H9 embryonic stem cells (H9, n=3) and primary hepatocytes (Heps, n=3) were used as controls. Gene expression is normalized to the housekeeping gene, UBC, and data is plotted as mean and SEM. W= WNT3A conditioned media, D=DAPT, SB = SB 431542, B4= BMP4, C=CHIR 99021, DZ= DBZ, EM = extrahepatic bile duct expansion media. # = $p \leq 0.05$, ## = $p \leq 0.01$, ### = $p \leq 0.001$ for the screening condition compared to EM. Experiments performed in collaboration with Ms. Simone Adams.

5.3.3 Transplantation of extrahepatic bile duct organoids into immunodeficient mice

Given our lack of success with inducing differentiation of our EHBD organoids *in vitro*, we attempted to assess if EHBD organoids have the capacity for engraftment in the mouse liver and if *in vivo* conditions may prompt a differentiation capacity that we were unable to discern *in vitro*. For this, we collaborated with Mr. Kourosh Saeb-Parsy in the Department of Surgery at Cambridge University, who transplanted EHBD organoids derived from a single common bile duct tissue donor, under the liver capsule of immunodeficient, 6-8 week old NOD-SCID IL2-gamma (NSG) mice. All animal studies were performed in accordance with UK Home Office regulations. In brief, dense cultures of EHBD organoids were collected by incubating with cell recovery solution at 4°C for 30 minutes. The organoids were then mechanically dissociated into medium sized clumps and washed twice with media. Organoid clumps were then re-suspended in 200-300uL of ice-cold Matrigel. Organoids and matrigel suspensions were then loaded into a pre-chilled syringe and kept on ice to prevent Matrigel polymerization.

Mice were anaesthetized with isoflurane, the abdominal cavity opened by a transverse laparotomy, and the liver exposed. The upper liver lobe was lifted using a sterile cotton swab and 30uL of matrigel was injected into the subcapsular space of the lower lobe. Approximately $2x10^6$ cells were injected per animal. These injections were performed on eight mice. Four mice were sacrificed at 2 weeks and the remaining four mice were sacrificed at three months. The livers were frozen in OCT and serially cryosectioned. Unfortunately, we were unable to find any human cells engrafted in the livers using a human specific nuclear antibody, Ku80 (**Figure 5.9**). We did observe in one of the livers, an area that appeared to be the site of injection, but no human cells remained (**Figure 5.10**).

These results suggest that EHBD organoids may not survive *in vivo* and do not have the capability to engraft in the mouse liver. However, it is also possible that we were unable to locate the cells, especially if only a small population remained engrafted. These experiments should be repeated before drawing definite conclusions.

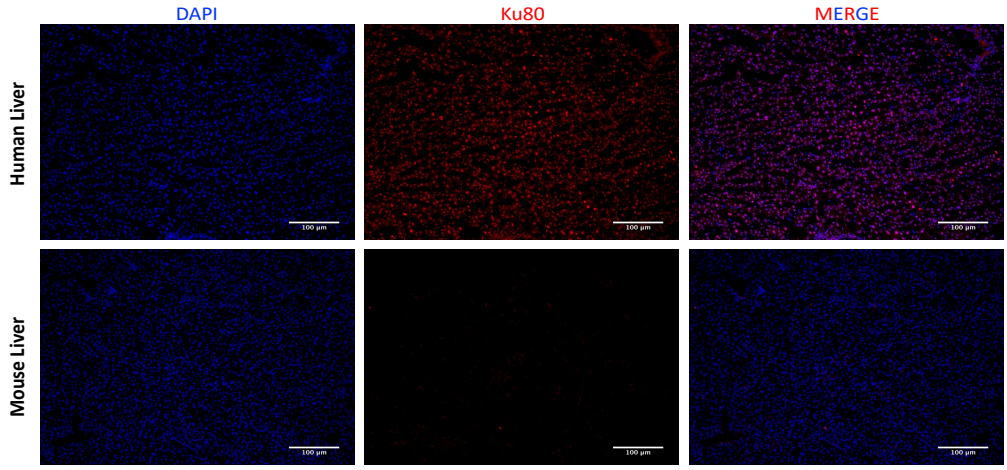


Figure 5.9 Immunofluorescence staining for a human-specific nuclear antibody, KU80, in human and mouse liver.

IF staining of human and mouse liver demonstrates that KU80 is specific for human cells and does not label mouse cells. Scale bar =100 um.

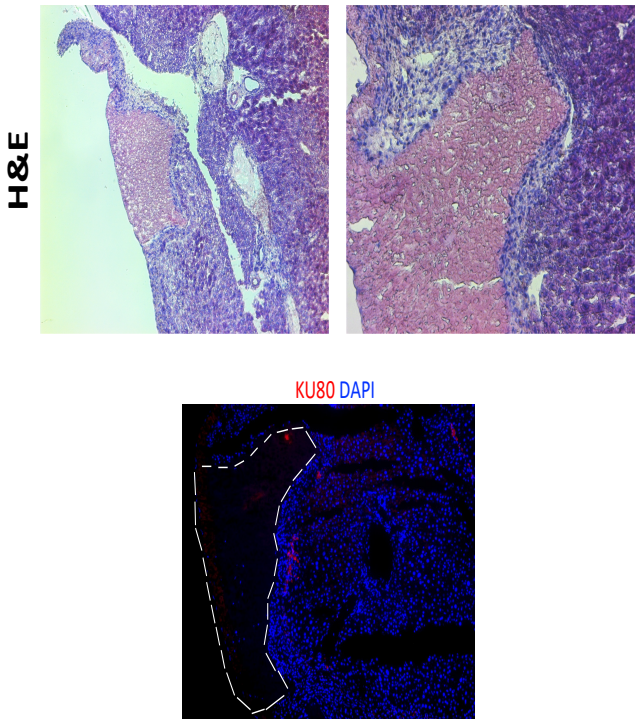


Figure 5.10 Haematoxylin and Eosin histology and Ku80 immunofluorescence staining of mouse liver injected with EHBD organoids under the liver capsule.

Images from one mouse liver that was injected with EHBD organoids, under the liver capsule, showing an area where it is likely the human cells were injected but did not remain engrafted. Top images show Haematoxylin and Eosin (H&E) staining of an acellular area and the bottom image shows IF analysis for KU80 indicating the absence of human cells engrafted under the mouse liver capsule.

5.4 Discussion

In this chapter, the differentiation capacity of EHBD and IHBD organoids was assessed. We took advantage of a previously published protocol that has been shown to increase the expression of hepatocyte specific markers in IHBD organoids. Using this protocol, it was discovered that only IHBD organoids up-regulate the expression of hepatocyte specific markers such as ALB, TTR, and CYP3A4, while EHBD organoids do not. Further, a wide-variety of compounds were tested on EHBD organoids to further evaluate their capacity for differentiation, but none of these results suggested a differentiation capacity towards a hepatocyte fate for EHBD organoids. Lastly, we attempted to transplant our EHBD organoids under the liver capsule of immunodeficient mice, but unfortunately, we were unable to locate any engrafted cells following transplantation.

We feel these results are significant in several ways. The first is that it continues to support our previous observations that IHBD and EHBD organoids are unique. Not only do these cells require different culture conditions and differ on the transcriptional level, but they also appear to have key differences in their capacity for differentiation. However, these results do raise several questions about the nature of the extrahepatic organoids.

If the EHBD organoids are not capable of differentiation towards a hepatocyte fate, can they be a stem/progenitor population? In all of our experiments, we only observed the EHBD organoids as having a phenotype consistent with that of biliary cells. However, our RNA-sequencing results demonstrated that despite expressing key biliary lineage markers such as KRT19/KRT7, these organoids down-regulate functional markers of biliary epithelial cells such as GGT1, HNF1B, MUC1.

Ultimately, we see two possible explanations for the origin of our EHBD organoids. First, it could be that, within the extrahepatic biliary tree, there is a rare population of cells which express adult stem cell markers and are selected for and expanded in our *in vitro* culture conditions. Alternatively, and more likely given the lack of adult stem cell marker expression in EHBD tissues, it could be that our *in vitro* culture conditions cause mature/differentiated extrahepatic biliary cells to undergo a type of “de-differentiation”. These cells down-regulate the expression of mature biliary markers and upregulate the expression of adSC markers, possibly in response to *in vitro* provided factors such as RSPO and WNT. In the intrahepatic compartment, it has been shown that WNT signalling is induced in response to tissue injury, which then activates intrahepatic biliary cells to

express LGR5 and proliferate.[98] It is possible that this mechanism in the intrahepatic biliary system, may also be at play in the extrahepatic system. Perhaps our *in vitro* culture system is mimicking a state of tissue injury. Within the intrahepatic compartment, which has a closer relationship both developmentally and spatially to the liver, biliary cells are capable of bipotential differentiation. This perhaps represents an embryologic memory, as intrahepatic cholangiocytes and hepatocytes share a common developmental progenitor, unlike the extrahepatic cholangiocytes. However, in the extrahepatic compartment, there is only one mature cell type. Despite sharing many characteristics with IHBD cells, EHBD cells, when exposed to WNT signalling, may also adopt an adSC and proliferative phenotype but cannot differentiate beyond a cholangiocyte fate. Perhaps then, these extrahepatic organoids may represent a biliary progenitor population which is only capable of biliary differentiation.

Lastly, it must be addressed that these results are contrary to several reports in the literature stating that EHBD derived cell-lines are capable of differentiation. In particular, work by Lola Reid and colleagues has suggested that EHBD derived cell lines from both CBD and GBD, are capable of differentiation towards a hepatocyte fate.[111,113] One explanation for the differences in our results is that we may have isolated and/or expanded different populations of cells from the biliary tree.

Interestingly, the bile duct cell lines isolated by Reid and colleagues are claimed to originate from the peribiliary glands (PBGs) and not the main epithelium. Their cells were isolated from bile ducts and gallbladders using enzymatic dissociation and EPCAM as a sorting marker. However, EPCAM is not a PBG specific marker and would isolate most, if not all, of the epithelial cells from the biliary epithelium and PBGs. Additionally, the GBD lacks PBGs.

It is possible, given our use of a mechanical dissociation technique, that we did not access this same population of cells, particularly in our CBD cell lines. However, given the larger overall percentage of epithelial cells compared to PBG cells, the ease at which epithelial cells were dissociated by scraping the mucosal surface with a scalpel, and the fact that the PBG cells are deeply embedded in underlying stromal tissue, it is likely that the cells isolated by Reid and colleagues also contains a large population of epithelial cells. Further, this would not explain the differences in results we observed in our GBD cell lines, as there are no PBGs in the GBD.

Perhaps then, the differentiation capability observed in their work may represent a response similar to what we observed in our organoids, where a very low-level of ALB/CYP3A4 expression was detected in stressed/dying cells. In fact, the fold change in ALB in their bile duct cell lines under differentiation conditions was only 2-fold higher than cells in their expansion conditions and was unfortunately not compared to primary hepatocytes or even IHBD cells. In contrast, we observed a 60-fold increase in Albumin expression in IHBD organoids. Although we did not directly test the differentiation media used by Reid and colleagues on our own EHBD organoids, we did screen all of the main individual factors present in their differentiation medium (i.e. Oncostatin-M, EGF, HGF, and Dexamethasone) and did not find any of these factors capable of inducing expression of hepatocyte markers in our cells. These two culture systems, and the cells they isolate, should be compared further in future studies in order to reconcile these discrepancies in results. However, what we are able to conclude, is that our results strongly suggest that, when compared to IHBD cells cultured as organoids, EHBD cells cultured as organoids do not have a capacity for differentiation towards a hepatocyte fate.

6 FUTURE DIRECTIONS AND CONCLUSION

6.1 Future directions

The discussions in Chapters 3, 4, and 5 outlined several future areas of research to further validate and expand upon the findings presented in the individual chapters of this dissertation. The limitations to the work were also discussed. In this chapter, we focus more broadly on the collective work of this dissertation and the future directions planned.

6.1.1 Extrahepatic biliary organoids

6.1.1.1 Further exploring the differentiation capacity of EHBD organoids

In this dissertation, we were able to comprehensively characterize our EHBD organoids and found that, when cultured in media conditions promoting canonical WNT signalling, biliary epithelial cells from the CBD, GBD, and PancD adopt a stem/progenitor-like phenotype with expression of adSC markers and long-term proliferative ability. However, it was also discovered that EHBD organoids do not appear to have a capacity for differentiation towards a hepatocyte-fate, whereas IHBD organoids are capable of upregulating expression of hepatocyte specific markers. These results are surprising, in some regards, given the *in vitro* characterizations and RNA-Seq data demonstrating the

EHBD organoids have some characteristics of a stem/progenitor phenotype. Therefore, it still remains unclear whether or not these EHBD organoids are truly a stem/progenitor population and if they have any capacity for differentiation. Here, we only focused on hepatocyte differentiation capacity and future work is still needed to assess for a differentiation capacity towards other cell lineages, especially towards either a mature cholangiocyte and/or pancreatic endocrine fate. To do this, we plan to take advantage of several recently published reports that have been described for differentiating intrapancreatic and intrahepatic biliary organoids towards a pancreatic endocrine and/or mature cholangiocyte fate.[95,149] We plan to screen these protocols on the EHBD organoids to assess for any differentiation capacity towards these cell fates using qPCR for lineage specific markers and functional profiling of the cells if we see any induction of these markers.

We also would like to explore the differences observed between the two culture systems developed in our lab to derive extrahepatic biliary organoids. As previously mentioned, biliary organoids were first derived in culture conditions containing DKK, a canonical WNT pathway inhibitor. It was found that the organoids cultured with DKK, maintained expression of functional biliary markers. However, in the system described in this dissertation, in which CHIR 99021 was used to promote canonical WNT signalling, the organoids instead expressed adSC markers and down-regulated functional biliary markers. We plan to further characterize the mechanistic role that WNT signalling may play in maintaining these two phenotypes and how these two culture systems differ through gene expression analyses.

Lastly, we plan to repeat the transplantation experiments already attempted in order to confirm the engraftment capacity of EHBD organoids and to confirm that our *in vitro* observations on the differentiation capacity remains true *in vivo*. To do so, we first plan to repeat the liver capsule injections of the EHBD organoids. However, to overcome the limitation previously encountered of finding engrafted cells, organoids will be fluorescently labelled before transplantation. Further, it would also be interesting to transplant organoids derived from both the IHBD and EHBDs into a mouse model of hepatocyte and/or cholangiocyte senescence.

Our collaborators recently reported that murine intrahepatic bile duct derived cells are capable of repopulating the liver and differentiating into hepatocytes in this model.[38] They are currently working to humanize this mouse model and it will serve as an invaluable *in vivo* tool.

6.1.2 Regional diversity in the biliary tree

6.1.2.1 Single cell RNA sequencing of intrahepatic and extrahepatic biliary tissues

Our RNA-sequencing experiments yielded a number of impactful results by characterizing the transcriptional profile of biliary tissues and *in vitro* derived organoids. However, one of the key limitations to our RNA-sequencing experiments was the contamination by non-biliary cell types in our primary tissue samples. This was particularly evident in PancD and IHBD tissue samples. Our lab has recently gained the ability to perform single cell RNA-sequencing and has been developing improved protocols for enzymatic dissociation of human tissues, including bile duct and liver tissue, that improves cell viability. This new expertise will allow us to circumvent some of the limitations encountered in our bulk RNA-sequencing data. We plan to perform single cell RNA-sequencing on EHBD and IHBD tissue samples in order to continue our efforts to profile and understand the regional diversity that exists in these tissues. In particular, by using single cell RNA-sequencing, we will not only be able to profile the epithelial biliary cells but can also profile other cell types in these tissues including the immune cells. This may ultimately help provide insight into the regionalization of autoimmune cholangiopathies within the biliary tree.

6.1.2 Regeneration in the extrahepatic biliary tree

Lastly, the results presented in this dissertation suggest that within the human extrahepatic biliary tree, cells are capable of adopting a stem/progenitor phenotype when cultured as *in vitro* organoids. This capability may be driven by canonical WNT signalling. However, these findings also raise questions on what functions these cells may play *in vivo* and whether or not these cells actually represent a regenerative population that plays a role in extrahepatic biliary tissue repair or homeostasis. Overall, it still is unknown what mechanisms drive regeneration in the extrahepatic biliary tree and this represents a large gap in knowledge. Whether or not a resident stem/progenitor cell population exists in these tissues *in vivo,* where the cells are located, or if there are cells that adopt a progenitor role during states of tissue injury is unknown. Therefore, we feel that this is an essential area in which future research is required and we propose several key experiments that

could be considered in the future to address this, some of which our lab has already begun working on.

6.1.2.1 Brief background on what is currently known on extrahepatic biliary regeneration

Many groups have suggested the existence of a stem/progenitor cell within the extrahepatic biliary tree and the data presented in this dissertation further supports the idea that human extrahepatic cholangiocytes are capable of adopting phenotypes of stem/progenitor cells *in vitro*.[53,54,110–117] In the large extrahepatic bile ducts and pancreatic ducts, it has been suggested that progenitor cells may reside in PBGs. PBGs are small epithelial/glandular structures that underlie the biliary epithelium. It was recently demonstrated that PBGs are directly connected to the biliary epithelium through small channels and that these cells express markers such as BrdU and KI67 and proliferate during bile duct injury.[53,54] However, Porte et al. (2012) also showed that, in human biliary disease, there is heterogeneity in which cellular compartment expresses proliferation markers during disease.[54] For example, in ischemic-type injury patients, where large portions of the epithelium are destroyed, the PBGs became the predominant source of proliferative cells.[54] Whereas in cholangitis patients, the epithelium appeared to be the strongest source of KI67 positive proliferating cells.[54] Further, in two mouse models of biliary injury: Bile Duct Ligation and RRV-virus infection, both the epithelium and peribiliary glands demonstrated BrdU uptake after injury.[53] These findings suggest that the regenerative compartment of the extrahepatic biliary system may not just be located in the PBGs. No study has definitively shown that PBG cells are capable of regenerating the epithelium of extrahepatic bile ducts or that progeny of PBG cells contribute to the bile duct epithelium. Furthermore, very little is known about the distribution of adSC markers like LGR5 and PROM1 in the PBG cells compared to the epithelium.

Also important to note, is that the GBD, another key organ in the extrahepatic biliary system, does not contain PBGs and instead consists of a highly folded mucosa. Potential stem/progenitor cells have been identified and cultured from the GBD epithelium by two different groups and also in this dissertation.[113,115–117] However, where these cells are

located *in vivo* and whether or not they actively participate in repair of the GBD epithelium is still not well known, as is true of the extrahepatic bile ducts.

Understanding what cellular compartment regenerates the extrahepatic biliary system is important as key human extrahepatic diseases such as PSC and Biliary Atresia may involve failure of regenerative mechanisms. Further, it has been shown that in human liver transplants, over 80-90% of all donor grafts show significant biliary epithelial cell loss before ever being transplanted and upwards of 40% of all patients receiving a liver transplant will go on to develop biliary complications.[129] These complications include non-anastomotic biliary strictures (NAS) in the large extrahepatic ducts, which are thought to result from a failure to regenerate the extrahepatic biliary epithelium after ischemic damage.[129,150] Interestingly, it was shown that NAS correlate strongly with the degree of histological damage to PBGs in biopsies taken of bile duct grafts prior to transplantation.[150]

With these ideas in mind, it is critical that future research works to identify the cellular compartment and also the cell population(s) responsible for regenerating the extrahepatic bile duct and gallbladder in states of tissue injury. Furthermore, we feel the results described in this dissertation provide a crucial foundation for addressing these questions.

6.1.2.2 Developing a reversible, acute injury model of the extrahepatic bile duct and gallbladder

One aspect that currently hinders further research into extrahepatic biliary regeneration is the lack of a reversible, acute extrahepatic injury model in mice. The most commonly used biliary injury model is bile duct ligation, but this procedure is irreversible and cannot be used to study the repair process following injury. Therefore, our lab has begun working to characterize a chemical bile duct injury model using 4,4'-methylene dianiline (DAPM) in mice. DAPM is a bile duct toxicant most well studied in Rats and the intrahepatic bile ducts, where it selectively injures biliary epithelial cells.[151] We have already begun work to validate its ability to cause injury in the extrahepatic biliary tree and hope this can serve as a model system to study biliary regeneration in the future.

6.1.2.3 Lineage tracing to identify stem/progenitor cells in the extrahepatic bile duct and gallbladder

Lineage tracing is a technique used to identify and track all of the progeny of a single cell.[152] It has been used to identify adSCs in various adult organs including intestine, stomach, and liver. Using the knowledge we gained from our human EHBD organoids, in particular, that the expression of PROM1 and LGR5 is upregulated in these organoids, we plan to attempt lineage tracing experiments in the extrahepatic biliary tree using these markers. Our lab currently has access to both LGR5 and PROM1 inducible reporter mouse lines and we plan to use these models to assess if these markers are expressed in the uninjured EHBDs of mice and also in DAPM injured EHBDs of mice. Ultimately, we feel that these experiments will allow us to build upon the findings presented in this dissertation and continue to further address these important open questions in the field.

6.2 Conclusions

According to the Cambridge guidelines for the award of PhD Degree, the candidate must demonstrate that his or her dissertation "represents a significant contribution to learning, for example through the discovery of new knowledge, the connection of previously unrelated facts, the development of new theory, or the revision of older views." To this end, this dissertation has accomplished these criteria through the following:

- Developing and extensively characterizing an *in vitro* organoid culture system capable of supporting cells from the human CBD, GBD, PancD, and IHBDs that express markers of adSCs.
- Demonstrating that organoids derived from human EHBD and IHBDs are unique and require different conditions for long-term proliferation.
- Describing for the first time the transcriptional profile of three human extrahepatic biliary tissue regions and uncovering that these tissues have key differences in expression and potential differences in biologic functions.
- Uncovering novel regional specific markers of human biliary tissues including SOX17, CDX2, and HOXB2.
- Revealing that EHBD organoids are transcriptionally unique from their tissue of origin, losing expression of mature biliary markers and upregulating expression of adSC markers *in vitro.*
- Presenting preliminary attempts to compare the transcriptional profile of IHBD and EHBD primary tissues.
- Discovering key differences in the differentiation capacity of EHBD and IHBD derived organoids towards a hepatocyte-fate.

Overall, this data has significantly furthered both the knowledge and tools available in the field. It has provided important insight to the understanding of both basic biliary physiology and regional diversity and also the field of biliary stem/progenitor cell organoids. These results may help to identify new targets for therapeutic development for cholangiopathies and regenerative medicine and has laid the groundwork for many future projects in the years to come.

7 REFERENCES

1. Amit M, Carpenter MK, Inokuma MS, et al. Clonally derived human embryonic stem cell lines maintain pluripotency and proliferative potential for prolonged periods of culture. Dev Biol 2000;227:271–278.
2. Hui H, Tang Y, Hu M, et al. Stem Cells: General Features and Characteristics. In: *Stem Cells in Clinic and Research*. InTech; 2011.
3. Health NI of. NIH Stem Cell Information. Stem Cell Inf [World Wide Web site] 2001. Available at: stemcells.nih.gov/info/2001report/chapter3.htm [Accessed January 2, 2018].
4. Wobus AM, Boheler KR. Embryonic Stem Cells : Prospects for Developmental Biology and Cell Therapy. Physiol Rev 85 2005;85:635–678.
5. Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. Nature 1981;292:154–156.
6. Thomson JA, Itskovitz-Eldor J, Shapiro SS, et al. Embryonic stem cell lines derived from human blastocysts. Science 1998;282:1145–7.
7. Vallier L. Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. J Cell Sci 2005;118:4495–4509.
8. Kallas A, Pook M, Trei A, et al. SOX2 Is Regulated Differently from NANOG and OCT4 in Human Embryonic Stem Cells during Early Differentiation Initiated with Sodium Butyrate. Stem Cells Int 2014;2014:298163.

9. Rodda DJ, Chew J-L, Lim L-H, et al. Transcriptional regulation of nanog by OCT4 and SOX2. J Biol Chem 2005;280:24731–7.

10. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 2006;126:663–76.

11. Takahashi K, Tanabe K, Ohnuki M, et al. Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. Cell 2007;131:861–872.

12. Shi Y, Inoue H, Wu JC, et al. Induced pluripotent stem cell technology: a decade of progress. Nat Rev Drug Discov 2017;16:115–130.

13. Touboul T, Hannan NRF, Corbineau S, et al. Generation of functional hepatocytes from human embryonic stem cells under chemically defined conditions that recapitulate liver development. Hepatology 2010;51:1754–65.

14. Hannan NRF, Segeritz C, Touboul T, et al. Production of hepatocyte like cells from human pluripotent stem cells. Nat Protoc 2013;8:430–437.

15. Gieseck RL, Hannan NRF, Bort R, et al. Maturation of induced pluripotent stem cell derived hepatocytes by 3D-culture. PLoS One 2014;9:e86372.

16. Grompe M. Tissue stem cells: new tools and functional diversity. Cell Stem Cell 2013;13:685–9.

17. Blanpain C, Horsley V, Fuchs E. Epithelial Stem Cells: Turning over New Leaves. Cell 2007;128:445–458.

18. Orkin SH, Zon LI. Hematopoiesis: An Evolving Paradigm for Stem Cell Biology. Cell 2008;132:631–644.

19. Snippert HJ, Haegebarth A, Kasper M, et al. Lgr6 Marks Stem Cells in the Hair Follicle That Generate All Cell Lineages of the Skin. Science (80-) 2010;327:1385–1389.

20. Bartfeld S, Bayram T, Wetering M van de, et al. In Vitro Expansion of Human Gastric Epithelial Stem Cells and Their Responses to Bacterial Infection. Gastroenterology 2015;148:126–136.e6.

21. Barker N, Es JH van, Kuipers J, et al. Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature 2007;449:1003–1007.

22. Cai C, Yu QC, Jiang W, et al. R-spondin1 is a novel hormone mediator for mammary stem cell self-renewal. Genes Dev 2014;28:2205–18.

23. Barker N, Rookmaaker MB, Kujala P, et al. Lgr5(+ve) stem/progenitor cells contribute to nephron formation during kidney development. Cell Rep 2012;2:540–52.

24. Karthaus WR, Iaquinta PJ, Drost J, et al. Identification of Multipotent Luminal Progenitor Cells in Human Prostate Organoid Cultures. Cell 2014;159:163–175.

25. Reya T, Clevers H. Wnt signalling in stem cells and cancer. Nature 2005;434:843–850.

26. Lau W de, Peng WC, Gros P, et al. The R-spondin/Lgr5/Rnf43 module: Regulator of Wnt signal strength. Genes Dev 2014;28:305–316.

27. Huelsken J. The Wnt signalling pathway. J Cell Sci 2002;115:3977–3978.

28. Koo B-K, Clevers H. Stem cells marked by the R-spondin receptor LGR5. Gastroenterology 2014;147:289–302.

29. Juza RM, Pauli EM. Clinical and surgical anatomy of the liver: A review for clinicians. Clin Anat 2014;27:764–769.

30. Abdel-Misih SRZ, Bloomston M. Liver Anatomy. Surg Clin North Am 2010;90:643–653.

31. Si-Tayeb K, Lemaigre FP, Duncan SA. Organogenesis and Development of the Liver. Dev Cell 2010;18:175–189.

32. Cox-North, Paula, Doorenbos, Ardith, Shannon, Sarah, Scott, John, Curtis JR. The Transition to End-of-Life Care in End-stage Liver Disease. J Hosp Palliat Nurs 2013;15:209–215.

33. NIDDK. Cirrhosis. Available at: http://digestive.niddk.nih.gov/ddiseases/pubs/Cirrhosis/index.aspx [Accessed September 2, 2014].

34. Anon. *OPTN/SRTR Annual Data Report: Liver*. 2012.

35. Zorn AM. Liver Development. StemBook 2008:1–26. Available at: http://www.stembook.org/.

36. Mao SA, Glorioso JM, Nyberg SL. Liver regeneration. Transl Res 2014;163:352–62.

37. Tanimizu N, Mitaka T. Re-evaluation of liver stem/progenitor cells. Organogenesis 2014;10:208–215.

38. Raven A, Lu WY, Man TY, et al. Cholangiocytes act as facultative liver stem cells during impaired hepatocyte regeneration. Nature 2017;547:350–354.

39. Font-Burgada. Hybrid Periportal Hepatocytes Regenerate the Injured Liver without Giving Rise to Cancer. Cell 2015;162:766–779.

40. Tanaka M, Itoh T, Tanimizu N, et al. Liver stem/progenitor cells: Their characteristics and regulatory mechanisms. J Biochem 2011;149:231–239.

41. Wang B, Zhao L, Fish M, et al. Self-renewing diploid Axin2+ cells fuel homeostatic renewal of the liver. Nature 2015;524:180–185.

42. Faber E. Similarities in the sequence of early histological changes induced in the liver of the rat by ethionine, 2-acetylamino fluorene, and 3′ -Methyl-4dimethylaminoazobenzene. Cancer Res 1956;16:142–8.

43. Lázaro CA, Rhim JA, Yamada Y, et al. Generation of hepatocytes from oval cell precursors in culture. Cancer Res 1998;58:5514–5522.

44. Lowes KN, Brennan BA, Yeoh GC, et al. Oval cell numbers in human chronic liver diseases are directly related to disease severity. Am J Pathol 1999;154:537–41.

45. Lu W-Y, Bird TG, Boulter L, et al. Hepatic progenitor cells of biliary origin with liver repopulation capacity. Nat Cell Biol 2015;17:971–83.

46. Bird TG, Forbes SJ. Two Fresh Streams to Fill the Liver's Hepatocyte Pool. Cell Stem Cell 2015;17:377–378.

47. Tabibian JH, Masyuk AI, Masyuk T V, et al. Physiology of cholangiocytes Terjung R, ed. Compr Physiol 2013;3:1–49.

48. Castaing D. Surgical anatomy of the biliary tract. HPB (Oxford) 2008;10:72–6.

49. Strazzabosco M, Fabris L. Functional anatomy of normal bile ducts. Anat Rec (Hoboken) 2008;291:653–60.

50. Hootman SR, Ondarza J. Overview of Pancreatic Duct Physiology and Pathophysiology. Digestion 1993;54:323–330.

51. Housset C, Chrétien Y, Debray D, et al. Functions of the Gallbladder. Compr Physiol 2016;6:1549–1577.

52. Cardinale V, Wang Y, Carpino G, et al. The biliary tree-a reservoir of multipotent stem cells. Nat Rev Gastroenterol Hepatol 2012;9:231–240.

53. DiPaola F, Shivakumar P, Pfister J, et al. Identification of intramural epithelial networks linked to peribiliary glands that express progenitor cell markers and proliferate after injury in mice. Hepatology 2013;58:1486–1496.

54. Sutton ME, op den Dries S, Koster MH, et al. Regeneration of human extrahepatic biliary epithelium: the peribiliary glands as progenitor cell compartment. Liver Int 2012;32:554–9.

55. Glaser SS, Gaudio E, Rao A, et al. Morphological and functional heterogeneity of the mouse intrahepatic biliary epithelium. Lab Investig 2009;89:456–469.

56. Glaser S, Francis H, Demorrow S, et al. Heterogeneity of the intrahepatic biliary epithelium. World J Gastroenterol 2006;12:3523–36.

57. Alpini G, Roberts S, Kuntz SM, et al. Morphological, molecular, and functional heterogeneity of cholangiocytes from normal rat liver. Gastroenterology 1996;110:1636–43.

58. Han Y, Glaser S, Meng F, et al. Recent advances in the morphological and functional heterogeneity of the biliary epithelium. Exp Biol Med 2013.

59. Alpini G, Glaser S, Robertson W, et al. Large but not small intrahepatic bile ducts are involved in secretin-regulated ductal bile secretion. Am J Physiol Liver Physiol 1997;272:G1064–G1074.

60. Ishii M, Vroman B, LaRusso NF. Isolation and morphologic characterization of bile duct epithelial cells from normal rat liver. Gastroenterology 1989.

61. Venter J, Francis H, Meng F, et al. Development and functional characterization of extrahepatic cholangiocyte lines from normal rats. Dig Liver Dis 2015.

62. Boyer JL. Bile formation and secretion. Compr Physiol 2013;3:1035–78.

63. Farina A, Dumonceau J-M, Lescuyer P. Proteomic Analysis of Human Bile and Potential Applications for Cancer Diagnosis. Expert Rev Proteomics 2009;6:285–301.

64. Chiang JYL. Bile acid metabolism and signaling. Compr Physiol 2013;3:1191–212.

65. Kalakonda A, John S. *Physiology, Bilirubin*. StatPearls Publishing; 2018.

66. Chignard N, Mergey M, Barbu V, et al. VPAC1 expression is regulated by FXR agonists in the human gallbladder epithelium. Hepatology 2005;42:549–557.

67. OpenStax. *Anatomy & Physiology*. 2013.

68. Longnecker D. Anatomy and Histology of the Pancreas. Pancreapedia Exocrine Pancreas … 2014:332–339.

69. Pandol SJ. *The Exocrine Pancreas*. 2010.

70. Li R, Zhang X, Yu L, et al. Characterization of Insulin-Immunoreactive Cells and Endocrine Cells Within the Duct System of the Adult Human Pancreas. Pancreas 2016;45:735–742.

71. Pallagi P, P&#233, Zolt&#225. The Physiology and Pathophysiology of Pancreatic Ductal Secretion: The Background for Clinicians. Pancreas 2015;44:1211–1233.

72. Park IS, Bendayan M. Development of the endocrine cells in the rat pancreatic and bile duct system. Histochem J 1993;25:807–20.

73. Lemaigre FP. Molecular mechanisms of biliary development. Prog Mol Biol Transl Sci 2010;97:103–126.

74. Zong Y, Stanger BZ. Molecular mechanisms of bile duct development. Int J Biochem Cell Biol 2011;43:257–264.

75. Lemaigre FP. Development of the biliary tract. Mech Dev 2003;120:81–87.

76. Strazzabosco M, Fabris L. Development of the bile ducts: Essentials for the clinical hepatologist. J Hepatol 2012;56:1159–1170.

77. Vakili K, Pomfret EA. Biliary Anatomy and Embryology. Surg Clin North Am 2008;88:1159–1174.

78. Spence JR, Lange AW, Lin SCJ, et al. Sox17 Regulates Organ Lineage Segregation of Ventral Foregut Progenitor Cells. Dev Cell 2009;17:62–74.

79. Gérard C, Tys J, Lemaigre FP. Gene regulatory networks in differentiation and direct reprogramming of hepatic cells. Semin Cell Dev Biol 2017;66:43–50.

80. Cordi S, Godard C, Saandi T, et al. Role of β-catenin in development of bile ducts. Differentiation 2016.

81. Prasanna LC. Accessory Pancreatic Duct Patterns and Their Clinical Implications. J Clin Diagnostic Res 2015;9:10–12.

82. Reichert M, Rustgi AK. Pancreatic ductal cells in development, regeneration, and neoplasia. J Clin Invest 2011;121:4572–4578.

83. Lazaridis KN, LaRusso NF. The Cholangiopathies. Mayo Clin Proc 2015;90:791–800.

84. Kelly DA, Davenport M. Current management of biliary atresia. Arch Dis Child 2007;92:1132–1135.

85. Chapman R, Cullen S. Etiopathogenesis of primary sclerosing cholangitis. World J Gastroenterol 2008;14:3350–3359.

86. Alhaque S, Themis M, Rashidi H. Three-dimensional cell culture: from evolution to revolution. 2018.

87. Clevers H. Modeling Development and Disease with Organoids. Cell 2016;165:1586–1597.

88. Sato T, Stange DE, Ferrante M, et al. Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. Gastroenterology 2011;141:1762–72.

89. Schmidt H, Zalyte R, Urnavicius L, et al. Vascularized and functional human liver from an iPSC-derived organ bud transplant. Nature 2015;516:435–438.

90. Kleinman HK, Martin GR. Matrigel: basement membrane matrix with biological activity. Semin Cancer Biol 2005;15:378–86.

91. Barker N, Clevers H. Lineage tracing in the intestinal epithelium. Curr Protoc Stem Cell Biol 2010:1–11.

92. Barker N. Adult intestinal stem cells: Critical drivers of epithelial homeostasis and regeneration. Nat Rev Mol Cell Biol 2014;15:19–33.

93. Yin X, Farin HF, Es JH van, et al. Niche-independent high-purity cultures of Lgr5+ intestinal stem cells and their progeny. Nat Methods 2014;11:106–12.

94. Lugli N, Kamileri I, Keogh A, et al. R-spondin 1 and noggin facilitate expansion of resident stem cells from non-damaged gallbladders. EMBO Rep 2016;17:769–79.

95. Loomans CJM, Williams Giuliani N, Balak J, et al. Expansion of Adult Human Pancreatic Tissue Yields Organoids Harboring Progenitor Cells with Endocrine Differentiation Potential. Stem Cell Reports 2018;10:712–724.

96. Huch M, Bonfanti P, Boj SF, et al. Unlimited in vitro expansion of adult bi-potent pancreas progenitors through the Lgr5/R-spondin axis. EMBO J 2013;32:2708–21.

97. Huch M, Gehart H, Boxtel R Van, et al. Long-term culture of genome-stable bipotent stem cells from adult human liver. Cell 2015:1–14.

98. Huch M, Dorrell C, Boj SF, et al. In vitro expansion of single Lgr5+ liver stem cells induced by Wnt-driven regeneration. Nature 2013;494:247–250.

99. Evarts R, Nagy P, Nakatsukasa H. In vivo differentiation of rat liver oval cells into hepatocytes. Cancer Res 1989;49:1541–1547.

100. Clouston AD, Powell EE, Walsh MJ, et al. Fibrosis correlates with a ductular reaction in hepatitis C: roles of impaired replication, progenitor cells and steatosis. Hepatology 2005;41:809–18.

101. Libbrecht L, Desmet V, Damme B Van, et al. Deep intralobular extension of human hepatic “progenitor cells” correlates with parenchymal inflammation in chronic viral hepatitis: can “progenitor cells” migrate? J Pathol 2000;192:373–8.

102. Lanthier N, Rubbia-Brandt L, Spahr L. Liver progenitor cells and therapeutic potential of stem cells in human chronic liver diseases. Acta Gastroenterol Belg 2013;76:3–9.

103. Yanger K, Knigin D, Zong Y, et al. Adult hepatocytes are generated by self-duplication rather than stem cell differentiation. Cell Stem Cell 2014;15:340–349.

104. Furuyama K, Kawaguchi Y, Akiyama H, et al. Continuous cell supply from a Sox9-expressing progenitor zone in adult liver, exocrine pancreas and intestine. Nat Genet 2011;43:34–41.

105. Shin S, Walton G, Aoki R, et al. Foxl1-Cre-marked adult hepatic progenitors have clonogenic and bilineage differentiation potential. Genes Dev 2011;25:1185–92.

106. Español-Suñer R, Carpentier R, Hul N Van, et al. Liver progenitor cells yield functional hepatocytes in response to chronic liver injury in mice. Gastroenterology 2012;143:1564–1575.e7.

107. Boulter, L. Lu, W.Y. Forbes SJ. Differentiation of progenitors in the liver: a matter of local choice. J Clin Invest 2013;123:1867–73.

108. Lorenzini S, Bird TG, Boulter L, et al. Characterisation of a stereotypical cellular and extracellular adult liver progenitor cell niche in rodents and diseased human liver. Gut 2011;59:645–654.

109. Tarlow BD, Pelz C, Naugler WE, et al. Bipotential Adult Liver Progenitors Are Derived from Chronically Injured Mature Hepatocytes. Cell Stem Cell 2014;15:605–618.

110. Irie T, Asahina K, Shimizu-Saito K, et al. Hepatic Progenitor Cells in the Mouse Extrahepatic Bile Duct after a Bile Duct Ligation. Stem Cells Dev 2007;16:979–988.

111. Cardinale V, Wang Y, Carpino G, et al. Multipotent stem/progenitor cells in human biliary tree give rise to hepatocytes, cholangiocytes, and pancreatic islets. Hepatology 2011;54:2159–72.

112. Wang Y, Lanzoni G, Carpino G, et al. Biliary tree stem cells, precursors to pancreatic committed progenitors: evidence for possible life-long pancreatic organogenesis. Stem Cells 2013;31:1966–79.

113. Carpino G, Cardinale V, Gentile R, et al. Evidence for multipotent endodermal stem/progenitor cell populations in human gallbladder. J Hepatol 2014;60:1194–202.

114. Semeraro R, Carpino G, Cardinale V, et al. Multipotent stem/progenitor cells in the human foetal biliary tree. J Hepatol 2012;57:987–94.

115. Manohar R, Komori J, Guzik L, et al. Identification and expansion of a unique stem cell population from adult mouse gallbladder. Hepatology 2011;54:1830–1841.

116. Lugli N, Kamileri I, Keogh A, et al. R-spondin 1 and noggin facilitate expansion of resident stem cells from non-damaged gallbladders. EMBO Rep 2016;17:769–79.

117. Manohar R, Li Y, Fohrer H, et al. Identification of a candidate stem cell in human gallbladder. Stem Cell Res 2015;14:258–269.

118. Engle DD, Corbo V, Jager M, et al. Organoid Models of Human and Mouse Ductal Pancreatic Cancer. Cell 2015;160:324–338.

119. Boj SF, Hwang C Il, Baker LA, et al. Organoid models of human and mouse ductal pancreatic cancer. Cell 2015;160:324–338.

120. Dorrell C, Tarlow B, Wang Y, et al. The organoid-initiating cells in mouse pancreas and liver are phenotypically and functionally similar. Stem Cell Res 2014;13:275–283.

121. Cardinale V, Gentile R, Ardizzoni A, et al. Human Biliary Tree Stem/Progenitor Cells immunomodulation: role of HGF.

122. Sampaziotis F, Justin AW, Tysoe OC, et al. Reconstruction of the mouse extrahepatic biliary tree using primary human extrahepatic cholangiocyte organoids. Nat Med 2017;23:954–963.

123. Broutier L, Andersson-Rolf A, Hindley CJ, et al. Culture and establishment of self-renewing human and mouse adult liver and pancreas 3D organoids and their genetic manipulation. Nat Protoc 2016;11:1724–1743.

124. Kraiczy J, Nayak KM, Howell KJ, et al. DNA methylation defines regional identity of human intestinal epithelial organoids and undergoes dynamic changes during development. Gut 2017:gutjnl-2017-314817.

125. Patro R, Duggal G, Love MI, et al. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 2017;14:417–419.

126. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Research 2016;4:1521.

127. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.

128. Eden E, Navon R, Steinfeld I, et al. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics 2009;10:48.

129. Karimian N, Op den Dries S, Porte RJ. The origin of biliary strictures after liver transplantation: is it the amount of epithelial injury or insufficient regeneration that counts? J Hepatol 2013;58:1065–7.

130. Forbester JL, Goulding D, Vallier L, et al. Interaction of Salmonella enterica Serovar Typhimurium with Intestinal Organoids Derived from Human Induced Pluripotent Stem Cells. Infect Immun 2015;83:2926–34.

131. Ferreira-Gonzalez S, Lu W-Y, Raven A, et al. Paracrine cellular senescence exacerbates biliary injury and impairs regeneration. Nat Commun 2018;9:1020.

132. Mu X, Pradere JP, Affò S, et al. Epithelial Transforming Growth Factor-β Signaling Does Not Contribute to Liver Fibrosis but Protects Mice from Cholangiocarcinoma. Gastroenterology 2016;150:720–733.

133. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 2009;10:57–63.

134. Kukurba KR, Montgomery SB. RNA Sequencing and Analysis. Cold Spring Harb Protoc 2016;2015:951–969.

135. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010;26:139–140.

136. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545–50.

137. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4:44–57.

138. Kampf C, Mardinoglu A, Fagerberg L, et al. Defining the human gallbladder proteome by transcriptomics and affinity proteomics. Proteomics 2014;14:2498–2507.

139. Fagerberg L, Hallström BM, Oksvold P, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol Cell Proteomics 2014;13:397–406.

140. Chapman MH, Tidswell R, Dooley JS, et al. Whole genome RNA expression profiling of endoscopic biliary brushings provides data suitable for biomarker discovery in cholangiocarcinoma. J Hepatol 2012;56:877–85.

141. Bertelli E, Bendayan M. Association between Endocrine Pancreas and Ductal System. More than an Epiphenomenon of Endocrine Differentiation and Development? J Histochem Cytochem 2005;53:1071–1086.

142. Muñoz J, Stange DE, Schepers AG, et al. The Lgr5 intestinal stem cell signature: Robust expression of proposed quiescent ′ +4′ cell markers. EMBO J 2012;31:3079–3091.

143. Sicklick JK, Li Y, Melhem A, et al. Hedgehog signaling maintains resident hepatic progenitors throughout life. 2006.

144. Hofmann JJ, Zovein AC, Koh H, et al. Jagged1 in the portal vein mesenchyme regulates intrahepatic bile duct development: insights into Alagille syndrome. Development 2010;137:4061–72.

145. Dorrell C, Schug J, Lin CF, et al. Transcriptomes of the major human pancreatic cell types. Diabetologia 2011;54:2832–44.

146. Xiao W, Hong H, Awadallah A, et al. Utilization of CDX2 Expression in Diagnosing Pancreatic Ductal Adenocarcinoma and Predicting Prognosis Batra SK, ed. PLoS One 2014;9:e86853.

147. Boulter L, Govaere O, Bird TG, et al. Macrophage-derived Wnt opposes Notch signaling to specify hepatic progenitor cell fate in chronic liver disease. Nat Med 2012;18:572–9.

148. Nagai F, Kato E, Tamura H. Oxidative stress induces GSTP1 and CYP3A4 expression in the human erythroleukemia cell line, K562. Biol Pharm Bull 2004;27:492–5.

149. Chen C, Jochems PGM, Salz L, et al. Bioengineered bile ducts recapitulate key cholangiocyte functions. Biofabrication 2018;10:034103.

150. op den Dries S, Westerkamp AC, Karimian N, et al. Injury to peribiliary glands and vascular plexus before liver transplantation predicts formation of non-anastomotic biliary strictures. J Hepatol 2014;60:1172–9.

151. Kanz MF, Gunasena GH, Kaphalia L, et al. A Minimally Toxic Dose of Methylene Dianiline Injures Biliary Epithelial Cells in Rats. 1998;426:414–426.

152. Kretzschmar K, Watt FM, Alvarez-Dolado M, et al. Lineage tracing. Cell 2012;148:33–45.

8 APPENDICES

Appendix I: List of Annex Files

Annex File 1 EHBD Tissue DGE.xlsx

List of genes differentially expressed between the three EHBD tissue regions.

Annex File 2 EHBD Tissue Specific Genes.xlsx

List of genes upregulated in one EHBD tissue region compared to both of the other two tissues. 256 genes were CBD specific, 419 genes were GBD specific, 967 genes were PancD specific.

Annex File 3 EHBD Tissue Gene Ontology Analyses.xlsx

Gene ontology analyses on the EHBD tissue specific gene lists and also genes differentially expressed between pairs of EHBD tissues.

Annex File 4 Genes Upregulated in GBD and CBD vs PancD Tissue.xlsx

List of 856 genes that were upregulated in both GBD and CBD tissues compared to PancD tissues.

Annex File 5 Genes Upregulated in GBD and CBD vs PancD Tissue Gene Ontology Analyses.xlsx

Gene ontology analyses on the 856 genes that were upregulated in both GBD and CBD tissues compared to PancD tissues.

Gene ontology analyses on 108 of these 856 genes that additionally were upregulated in GBD tissue compared to CBD tissue.

Annex File 6 EHBD Tissue vs Organoids DGE.xlsx

List of genes differentially expressed between EHBD tissues and organoids.

Annex File 7 EHBD Tissue vs Organoids Gene Ontology

Gene ontology analyses on the top 200 differentially expressed genes between each of the three EHBD tissues compared to their corresponding *in vitro* organoids.

Annex File 8 EHBD Organoids vs Organoids DGE.xlsx

List of genes differentially expressed between the three EHBD organoid types.

Annex File 9 Genes Upregulated in both EHBD Organoids and Tissues.xlsx

List of genes that were upregulated in CBD, GBD, or PancD tissues which were also upregulated in the corresponding *in vitro* Organoids when compared to the other two regions.

Annex File 10 Comparing EHBD Organoids to ISC Signature.xlsx

List of 493 genes making up the ISC signature.

List of genes that were upregulated in CBD, GBD, or PancD organoids compared to their tissue of origin which were represented in the list of 493 ISC signature genes.

Gene ontology analyses of 135 ISC signature genes which were upregulated in at least two out of the three EHBD organoid types compared to their tissue of origin.

Annex File 11 EHBD vs IHBD Organoids DGE.xlsx

List of genes differentially expressed between IHBD_ex, IHBD_Huch, and EHBD organoids.

List of genes upregulated in IHBD_ex organoids compared to both IHBD_Huch and EHBD organoids.

List of genes upregulated in both IHBD_Huch and EHBD compared to IHBD_ex organoids (a.k.a genes downregulated in IHBD_ex organoids).

Annex File 12 IHBD_Huch vs EHBD Organoids Gene Ontology Analyses.xlsx

Gene ontology analyses of genes differentially expressed between IHBD_Huch organoids and EHBD organoids

Annex File 13 IHBD vs EHBD Tissues DGE.xlsx

List of genes differentially expressed between IHBD tissue and CBD, GBD, or PancD tissues.

List of 2,595 genes that were upregulated in IHBD tissue when compared to CBD, GBD, and PancD tissues.

Gene ontology analyses on the 2,595 genes that were upregulated in IHBD tissue compared to CBD, GBD, and PancD tissues.

Appendix II: *DeSeq2* Code used for RNA-Seq analyses

Contents

In this script, I perform RNA-seq DGE analysis on primary tissue and *in vitro* cultured organoid cell lines from human common bile duct (CBD), gallbladder (GBD), pancreatic duct (PancD), and intrahepatic bile ducts (IHBD). All biological groups have at least 3 biological replicates per conditions.

Loading Libraries

First, I load the required R packages and set the desired working directory.

```
library("DESeq2")
library("tximport")
library("readr")
library("tximportData")
library("org.Hs.eg.db")
library("AnnotationDbi")
library("pheatmap")
library("RColorBrewer")
library("biomaRt")
library("pheatmap")
library("ggplot2")
dir <- "/users/rimlandca/Desktop/R files"
```

```
setwd(dir)
library(knitr)
```

I. Preliminary data exploration

A. Loading Data (all samples)

Next I load files containing sample group information (samples) and previously generated SALMON transcipt abundance files for each sample (files).

```
samples_ALL <- read.table(file.path(dir, "samples_v3_noliver.txt"), header = TRUE)
files_ALL <- file.path(dir, "salmon", paste("salmon_", samples_ALL$run, ".txt",
    sep = ""))
names(files_ALL) <- samples_ALL$run
rownames(samples_ALL) <- samples_ALL$run
samples_ALL$group <- samples_ALL$group
samples_ALL$dge <- samples_ALL$dge
samples_ALL$type <- samples_ALL$type
samples_ALL$techrep <- samples_ALL$techrep
samples_ALL$libbatch <- samples_ALL$libbatch
samples_ALL$seqyear <- samples_ALL$seqyear
samples_ALL$name <- samples_ALL$name
samples_ALL$sampletype <- samples_ALL$sampletype
```

B. Creating a DeSeq Data Set (all samples)

Next, I create a DeSeq data set using the tximport function.

```
load("/Users/rimlandca/Desktop/R files/tx2gene_GRCh38.80.RData")
txi_ALL <- tximport(files_ALL, type = "salmon", tx2gene = tx2gene)
ddsTxi_ALL <- DESeqDataSetFromTximport(txi_ALL, colData = samples_ALL, design = ~type)
```

C. Assessing for batch effects

Before continuing with further analyses, we must assess for any batch effects present in the data. Due to sample availability timing, we had to prepare libraries and sequence the samples in two batchs (Library Preparation 1 vs 2). Additionally some of the samples originally sequenced in the first batch had to be re-sequenced due to low read depth therefore we also have the potential for a batch effect by year of sequencing (Sequencing Year A vs B). To help account for variability between batches and serve as an internal control, 4 samples from the first library batch were re-prepped and sequenced with the second batch.

```
# Filter out genes with 0 counts across all samples
ddsTxi_batch <- ddsTxi_ALL[rowSums(counts(ddsTxi_ALL)) > 0, ]
nrow(ddsTxi_batch)
```

```
## [1] 54622
```

```
# Subset for just protein coding genes
collinfocoding <- read_csv("collinfocoding_nodups.csv", col_names = TRUE)
ddsTxi_batch_proteincoding <- subset(ddsTxi_batch, rownames(ddsTxi_batch) %in%
    collinfocoding$ID)
nrow(ddsTxi_batch_proteincoding)
```

```
## [1] 19376
```

```
# normalize counts with varaince stabilizing transformation
vsd_ddstxi_batch <- vst(ddsTxi_batch_proteincoding)
```

Plot PCA plots

All Samples by Type and Batch

```
pcadata_batch_All_type <- plotPCA(vsd_ddstxi_batch, intgroup = c("type"), ntop = 19376,
    returnData = TRUE)
percentVar_batch_All_type <- round(100 * attr(pcadata_batch_All_type, "percentVar"))
ggplot(pcadata_batch_All_type, aes(PC1, PC2, color = type)) + geom_point(size = 6) +
    xlab(paste0("PC1: ", percentVar_batch_All_type[1], "% variance")) + ylab(paste0("PC2: ",
    percentVar_batch_All_type[2], "% variance"))
```

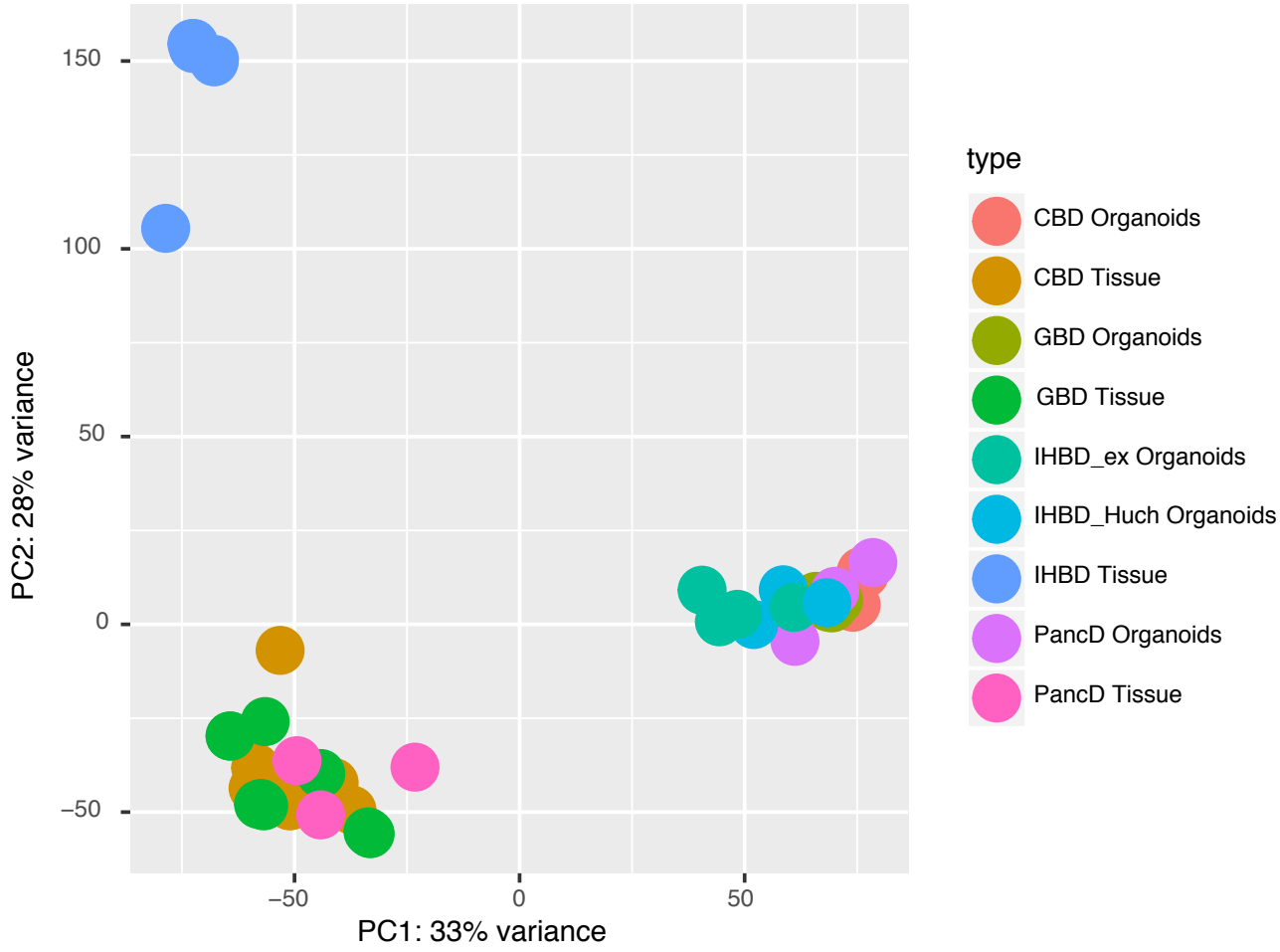


Figure 1: PCA of All Samples by Type

```
pcadata_batch_All_batch <- plotPCA(vsd_ddstxi_batch, intgroup = c("libbatch",
    "seqyear"), ntop = 19376, returnData = TRUE)
percentVar_batch_All_batch <- round(100 * attr(pcadata_batch_All_batch, "percentVar"))
ggplot(pcadata_batch_All_batch, aes(PC1, PC2, color = group)) + geom_point(size = 6) +
    xlab(paste0("PC1: ", percentVar_batch_All_batch[1], "% variance")) + ylab(paste0("PC2: ",
    percentVar_batch_All_batch[2], "% variance"))
```

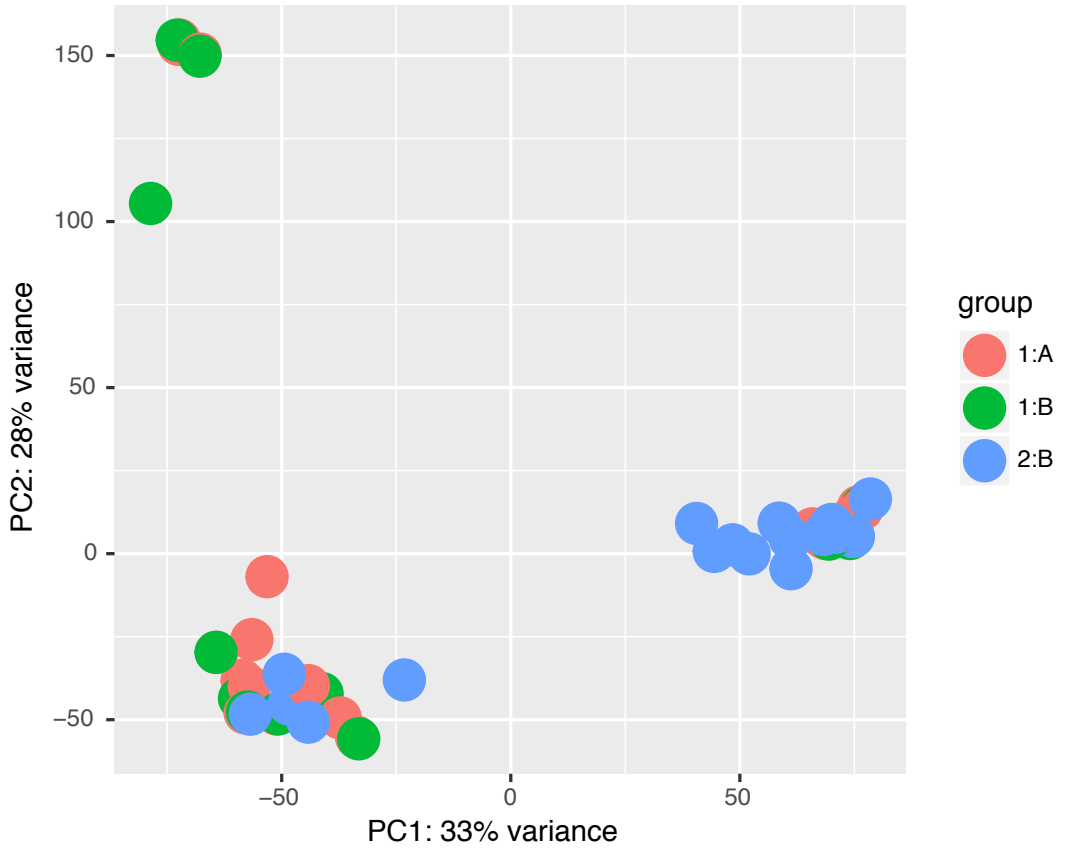


Figure 2: PCA of All Samples by Batch

Batch Effect Control Samples by Type and Batch

```
vsd_ddstxi_batchctrls <- vsd_ddstxi_batch[, vsd_ddstxi_batch$name == "GBD11_c" |
    vsd_ddstxi_batch$name == "GBD8_t" | vsd_ddstxi_batch$name == "PD5_c" | vsd_ddstxi_batch$name ==
    "PD5_t"]
pcadata_batchctrls_type <- plotPCA(vsd_ddstxi_batchctrls, intgroup = c("type"),
    ntop = 19376, returnData = TRUE)
percentVar_batchctrls_type <- round(100 * attr(pcadata_batchctrls_type, "percentVar"))
ggplot(pcadata_batchctrls_type, aes(PC1, PC2, color = type)) + geom_point(size = 6) +
    xlab(paste0("PC1: ", percentVar_batchctrls_type[1], "% variance")) + ylab(paste0("PC2: ",
    percentVar_batchctrls_type[2], "% variance"))
```

```
pcadata_batchctrls_batch <- plotPCA(vsd_ddstxi_batchctrls, intgroup = c("libbatch",
    "seqyear"), ntop = 19376, returnData = TRUE)
percentVar_batchctrls_batch <- round(100 * attr(pcadata_batchctrls_batch, "percentVar"))
ggplot(pcadata_batchctrls_batch, aes(PC1, PC2, color = group)) + geom_point(size = 6) +
    xlab(paste0("PC1: ", percentVar_batchctrls_batch[1], "% variance")) + ylab(paste0("PC2: ",
    percentVar_batchctrls_batch[2], "% variance"))
```

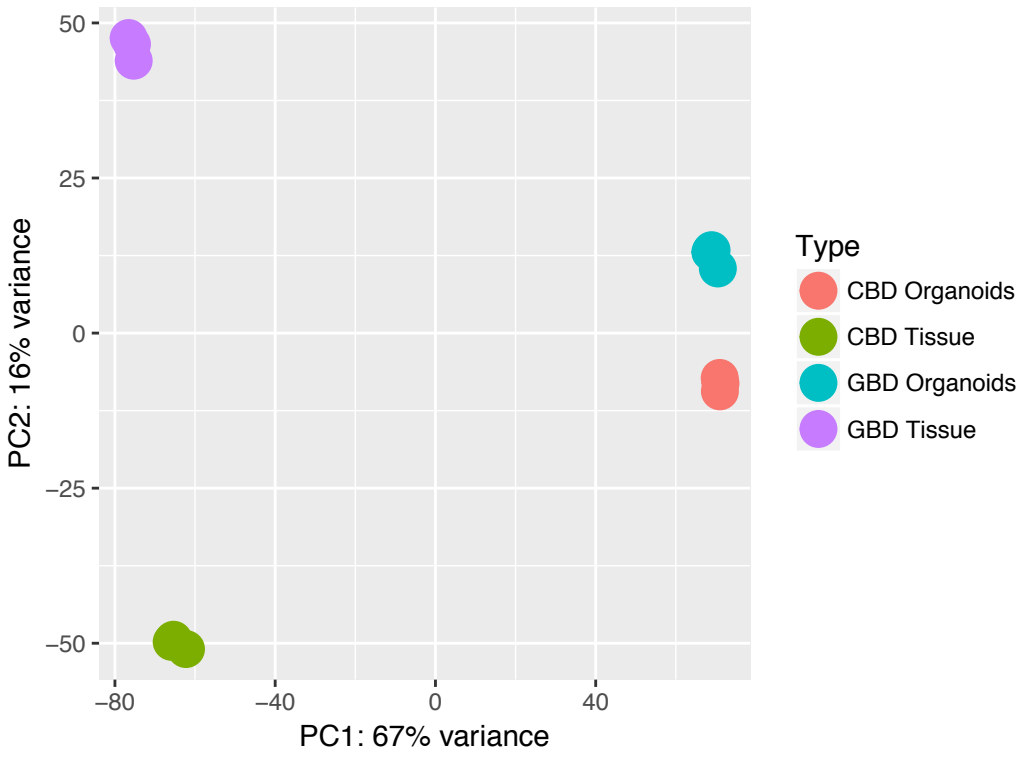


Figure 3: PCA of Batch Controls by Type

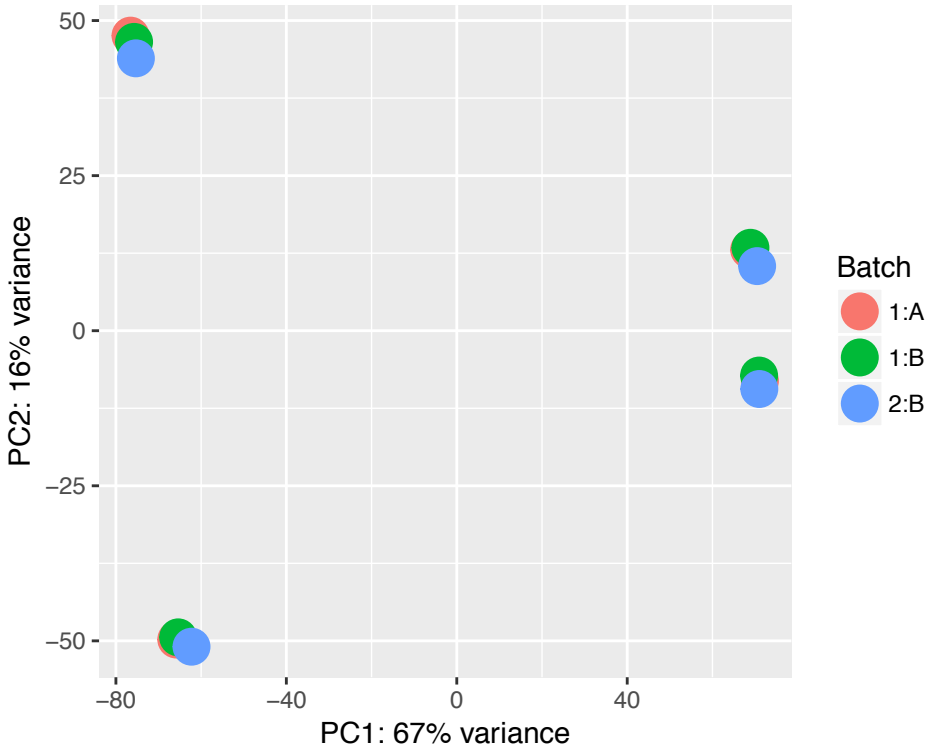


Figure 4: PCA of Batch Controls by Batch

D. Collapsing Technical Replicates, pre-filter, and subset for just protein coding genes

As we could not observe any significant batch effects in the data by sequencing year or libary preparation date, we collapse these technical replicates into single samples. We also filter for only genes which have greater than 0 counts for all samples and filter for just protein coding genes.

```
# Collapse replicates
ddsColl_ALL <- collapseReplicates(ddsTxi_ALL, samples_ALL$name)

## Pre-Filter
nrow(ddsColl_ALL)
```

```
## [1] 61692
```

```
ddsColl_ALL <- ddsColl_ALL[rowSums(counts(ddsColl_ALL)) > 0, ]
nrow(ddsColl_ALL)
```

```
## [1] 54622
```

```
# Subset for protein coding genes (protein coding gene data from biomart
# website)
ddsCollcoding_ALL <- subset(ddsColl_ALL, rownames(ddsColl_ALL) %in% collinfocoding$ID)
nrow(ddsCollcoding_ALL)
```

```
## [1] 19376
```

E. Principal component analyses

With technical replicates collapsed and pre-filtering performed, we now perform preliminary PCA analyses on the data set in order to assess sample relationships as well as identify any outliers.

We do find that one IHBD_ex sample appears to be an outlier. This sample was sequenced at passage 3 and was still proliferative while the others were at passage 5, given this known difference we exclude this sample from further analysis.

We also note that IHBD Tissue clusters very far from EHBD Tissues. As the IHBD tissue samples were from frozen, enzymatically dissociated liver tissue, and possibly contains contaminating cell types other than biliary cells, we will exclude these samples from analyses for now and focus on EHBD tissues and the EHBD/IHBD organoids only in section II. Exploratory analyses will be performed on the IHBD tissue in section III.

PCA of All Samples with Replicates Collapsed

```
vsd_coll_ALL <- vst(ddsCollcoding_ALL)
plotPCA(vsd_coll_ALL, "type", ntop = 19376)
```

PCA of IHBD Organoids

```
vsd_IHBDorganoids <- vsd_coll_ALL[, vsd_coll_ALL$group == "cells" & vsd_coll_ALL$EHBD_IHBD ==
    "IHBD"]
pcadata_IHBDorganoids <- plotPCA(vsd_IHBDorganoids, intgroup = c("type"), ntop = 19376,
    returnData = TRUE)
percentVar_IHBD <- round(100 * attr(pcadata_IHBDorganoids, "percentVar"))
ggplot(pcadata_IHBDorganoids, aes(PC1, PC2, color = type)) + geom_point(size = 6) +
    xlab(paste0("PC1: ", percentVar_IHBD[1], "% variance")) + ylab(paste0("PC2: ",
    percentVar_IHBD[2], "% variance"))
```

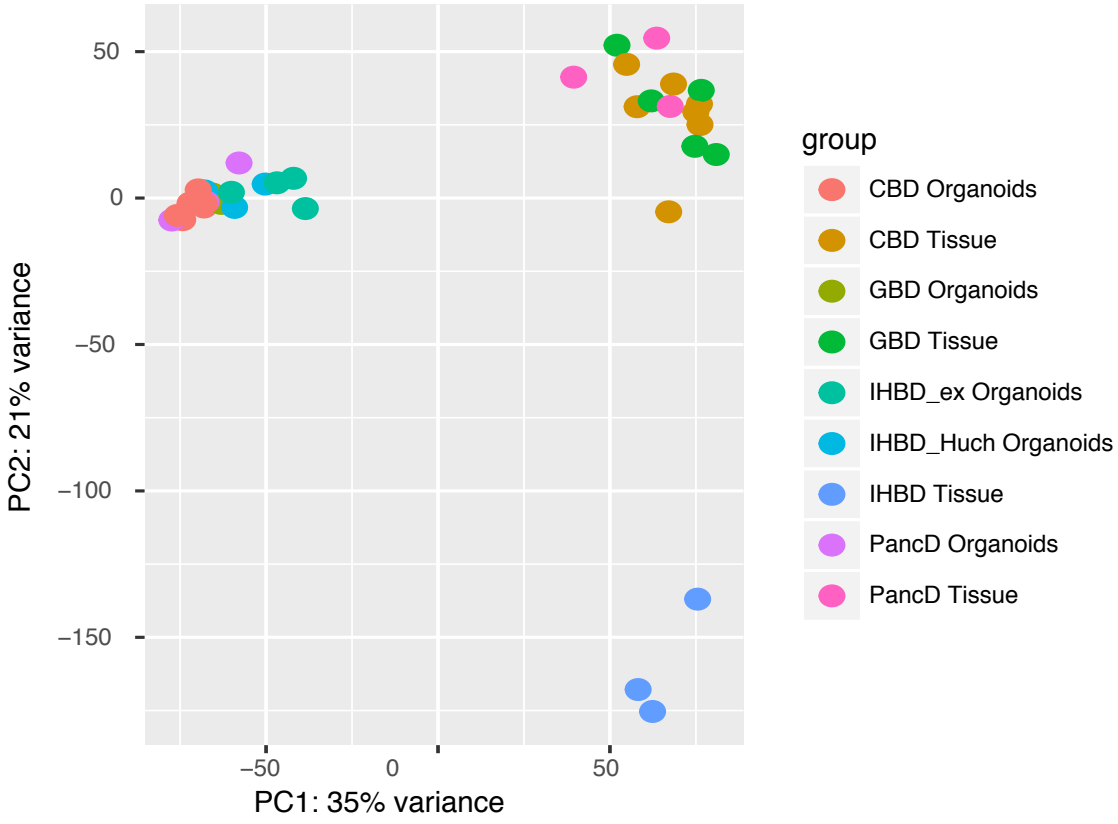


Figure 5: PCA of All Samples with Technical Replicates Collapsed

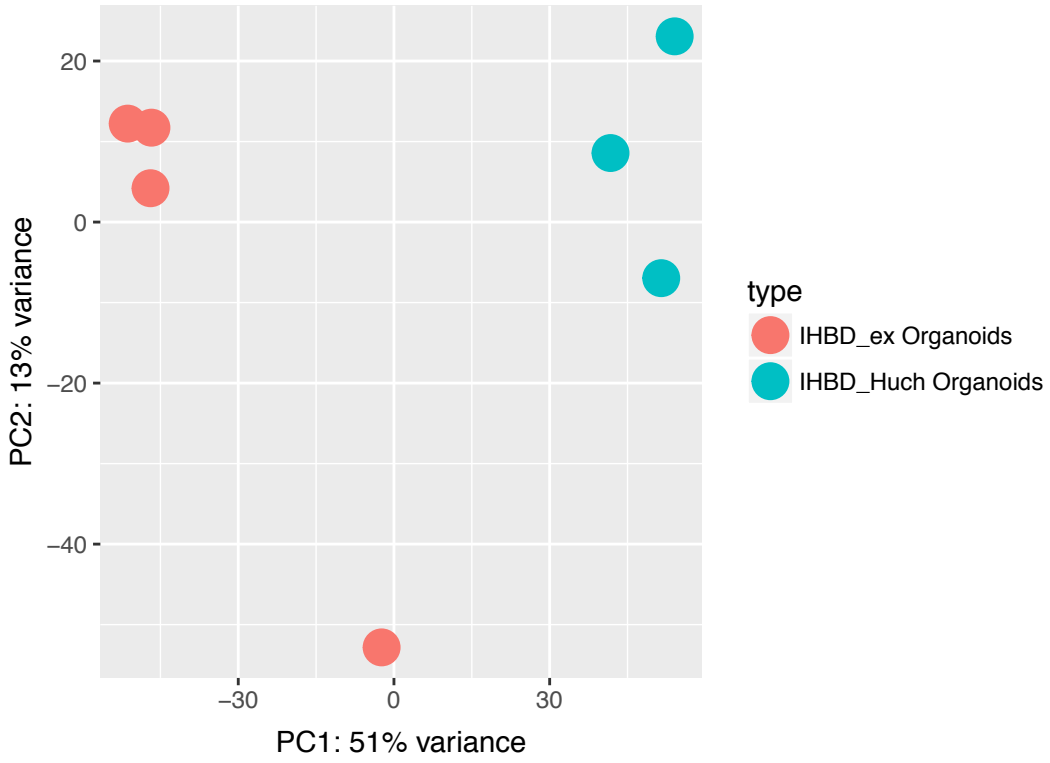


Figure 6: PCA of IHBD Organoids

II. Differential expression analyses

A. Load data for only the subset of samples

(excluding IHBD tissue samples and the outlier IHBD_ex sample)

```
# Loading Data without IHBD tissue samples or IHBD_ex outlier
samples <- read.table(file.path(dir, "samples_paper.txt"), header = TRUE)
files <- file.path(dir, "salmonpaper", paste("salmon_", samples$run, ".txt",
    sep = ""))
names(files) <- samples$run
rownames(samples) <- samples$run
samples$group <- samples$group
samples$dge <- samples$dge
samples$type <- samples$type
samples$techrep <- samples$techrep
samples$libbatch <- samples$libbatch
samples$seqyear <- samples$seqyear
samples$name <- samples$name
samples$sampletype <- samples$sampletype

# Create a DeSeq Matrix
txi <- tximport(files, type = "salmon", tx2gene = tx2gene)
ddsTxi <- DESeqDataSetFromTximport(txi, colData = samples, design = ~type)

# Collapsing Technical Replicates
ddsColl <- collapseReplicates(ddsTxi, samples$name)

# Pre-Filter
nrow(ddsColl)
```

```
## [1] 61692
```

```
ddsColl <- ddsColl[rowSums(counts(ddsColl)) > 0, ]
nrow(ddsColl)
```

```
## [1] 53803
```

```
# Subset protein coding genes
ddsCollcoding <- subset(ddsColl, rownames(ddsColl) %in% collinfocoding$ID)
nrow(ddsCollcoding)
```

```
## [1] 19327
```

B. Make PCA Plots

PCA Plot of EHBD Tissue Samples

```
vsd_coll <- vst(ddsCollcoding)
# Just EHBD Tissues
vsd_coll_Tissue <- vsd_coll[, vsd_coll$group != "cells"]
pcadata_tissues <- plotPCA(vsd_coll_Tissue, intgroup = c("type"), ntop = 5000,
    returnData = TRUE)
percentVar_tissues <- round(100 * attr(pcadata_tissues, "percentVar"))
```

```
ggplot(pcadata_tissues, aes(PC1, PC2, color = type)) + geom_point(size = 6,
    shape = 17) + xlab(paste0("PC1: ", percentVar_tissues[1], "% variance")) +
    ylab(paste0("PC2: ", percentVar_tissues[2], "% variance"))
```

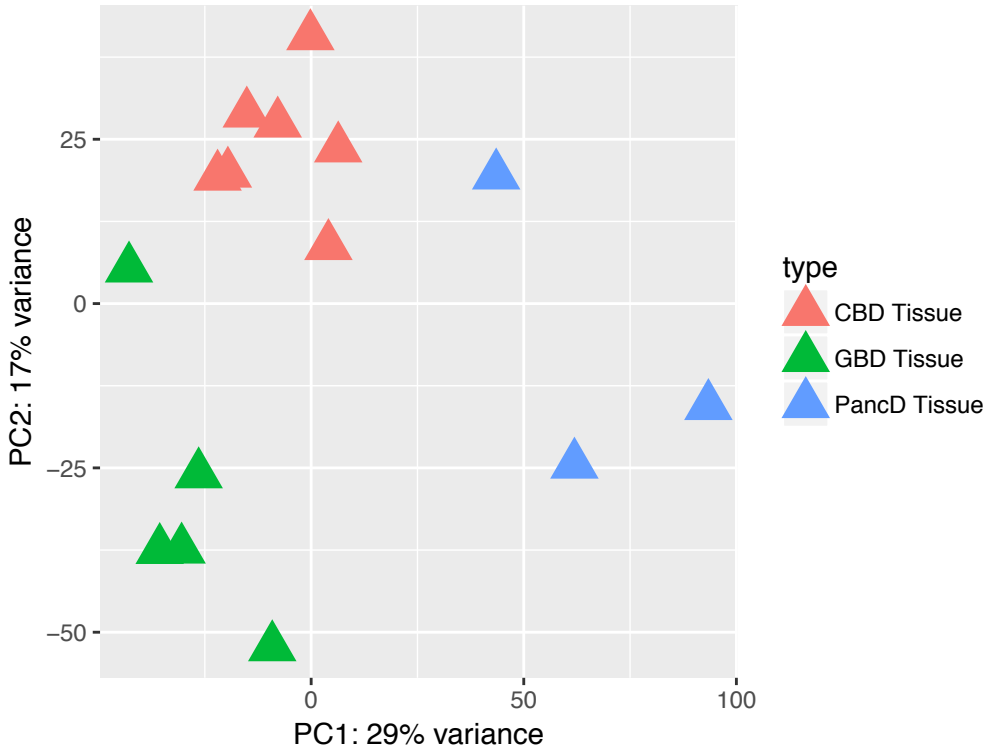


Figure 7: PCA of EHBD Tissues

PCA Plot of EHBD Tissue and Organoids

```
vsd_coll_EHBD <- vsd_coll[, vsd_coll$EHBD_IHBD != "IHBD"]
pcadata <- plotPCA(vsd_coll_EHBD, intgroup = c("group", "origin"), ntop = 5000,
    returnData = TRUE)
percentVar <- round(100 * attr(pcadata, "percentVar"))
ggplot(pcadata, aes(PC1, PC2, color = origin, shape = group.1)) + geom_point(size = 6) +
    xlab(paste0("PC1: ", percentVar[1], "% variance")) + ylab(paste0("PC2: ",
    percentVar[2], "% variance"))
```

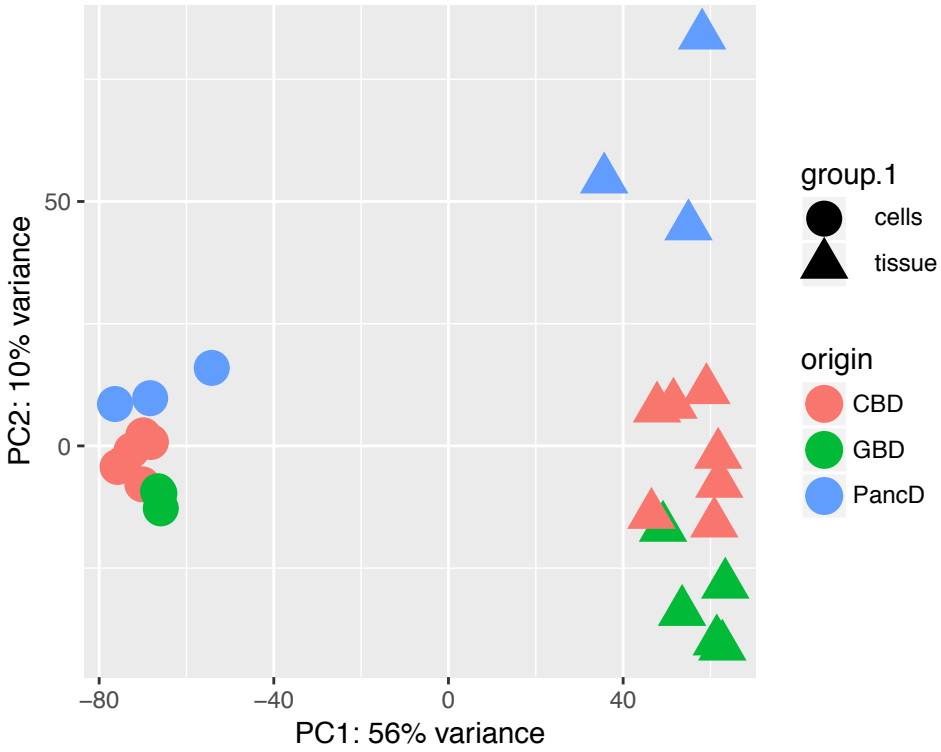


Figure 8: PCA of EHBD Tissue and Organoids

PCA Plot of EHBD and IHBD Organoids

```
vsd_coll_cells <- vsd_coll[, vsd_coll$group != "tissue"]
pcadata_cells <- plotPCA(vsd_coll_cells, intgroup = c("type"), ntop = 5000,
    returnData = TRUE)
percentVar_cells <- round(100 * attr(pcadata_cells, "percentVar"))
ggplot(pcadata_cells, aes(PC1, PC2, color = type)) + geom_point(size = 6) +
    xlab(paste0("PC1: ", percentVar_cells[1], "% variance")) + ylab(paste0("PC2: ",
    percentVar_cells[2], "% variance"))
```

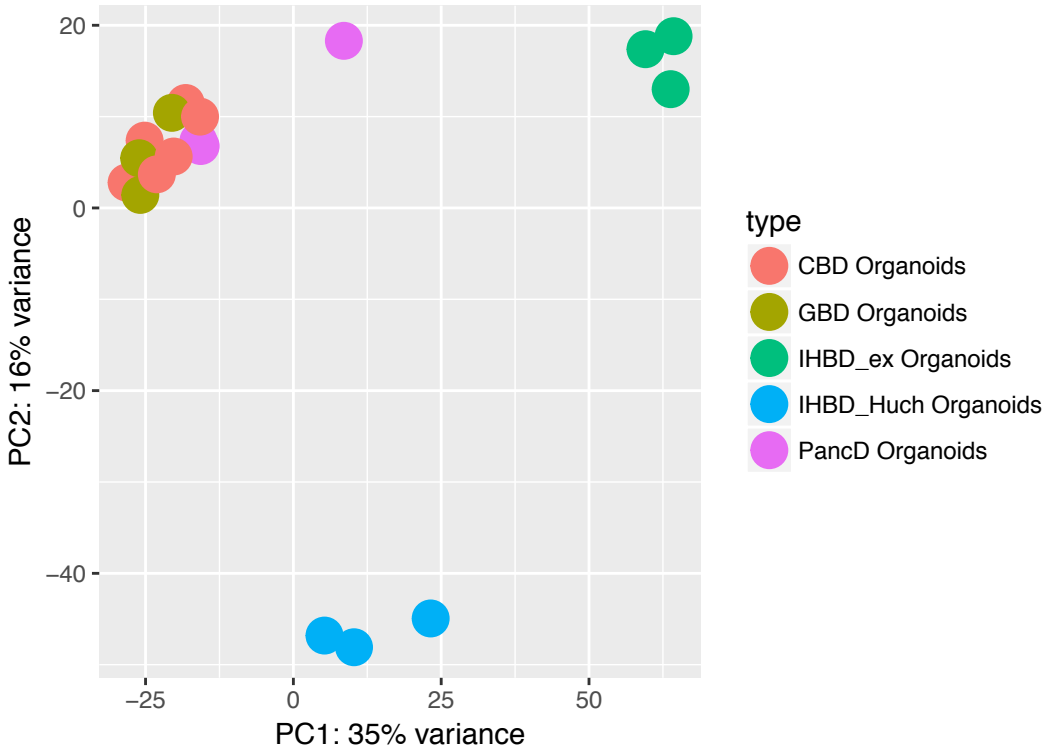


Figure 9: PCA of EHBD Tissue and Organoids

C. Export VST counts for making heatmaps

```
mat_vsdColl <- assay(vsd_coll)
mat_vsdColl.df <- as.data.frame(mat_vsdColl)
mat_vsdColl.df$ID <- rownames(mat_vsdColl.df)
mat_vsdColl.df <- merge(collinfocoding, mat_vsdColl.df, by = c("ID"), all.y = TRUE)
write_delim(mat_vsdColl.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/vsd_normalized_allsamples.csv")
```

D. Pearson Correlation Matrices for Organoid Types and Tissue Types

Pearson Correlation Matrix EHBD Tissues

```
corrs_EHBDtissues <- cor(assay(vsd_coll_Tissue), method = "pearson")
corr.dists_EHBDtissues <- as.dist(1 - corrs_EHBDtissues)
library("pheatmap")
colors <- colorRampPalette(c("red", "white", "blue"))(99)
pheatmap(corrs_EHBDtissues, breaks = seq(from = -1, to = 1, length = 100), clustering_distance_rows =
corr.dists_EHBDtissues, clustering_distance_cols = corr.dists_EHBDtissues, col = colors)
```

```
library(RColorBrewer)
diag(corrs_EHBDtissues) <- NA
colors <- colorRampPalette(brewer.pal(9, "Blues"))(99)
pheatmap(corrs_EHBDtissues, clustering_distance_rows = corr.dists_EHBDtissues,
    clustering_distance_cols = corr.dists_EHBDtissues, col = colors)
```

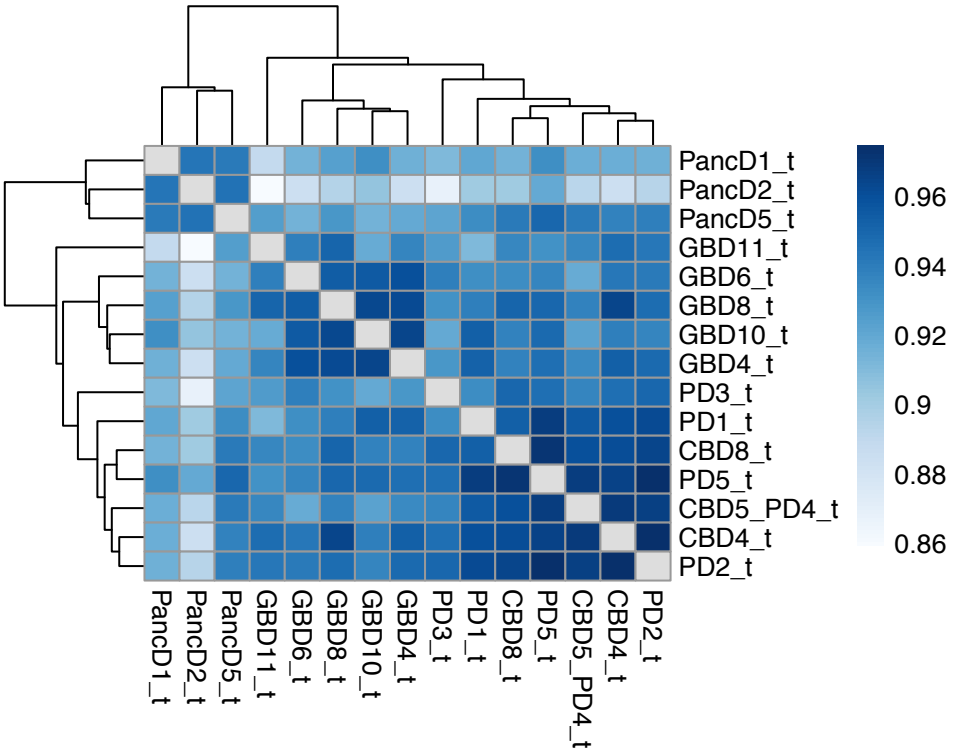


Figure 11: Correlation Matrix EHBD Tissues

Pearson correlation matrix EHBD and IHBD Organoids

```
corrs_Organoids <- cor(assay(vsd_coll_cells), method = "pearson")
corr.dists_Organoids <- as.dist(1 - corrs_Organoids)
library("pheatmap")
colors <- colorRampPalette(c("red", "white", "blue"))(99)
pheatmap(corrs_Organoids, breaks = seq(from = -1, to = 1, length = 100), clustering_distance_rows =
corr. dists_Organoids, clustering_distance_cols = corr.dists_Organoids, col = colors)
```

```
library(RColorBrewer)
diag(corrs_Organoids) <- NA
colors <- colorRampPalette(brewer.pal(9, "Blues"))(99)
pheatmap(corrs_Organoids, clustering_distance_rows = corr.dists_Organoids, clustering_distance_cols =
corr. dists_Organoids, col = colors)
```

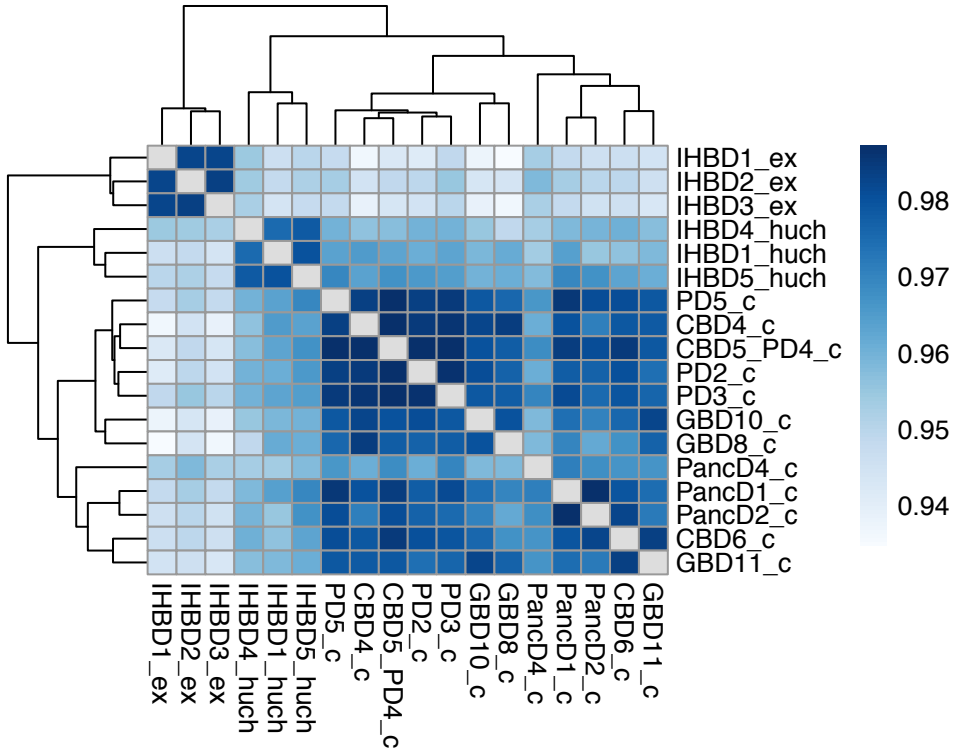


Figure 13: Correlation Matrix EHBD vs IHBD Organoids

E. Perform differential gene expression analyses

```
# Run DeSeq2
dds_DGE <- DESeq(ddsCollcoding, betaPrior = TRUE)

# Get base mean expression levels for each individual group
baseMeanPerLvl <- sapply(levels(dds_DGE$type), function(lvl) rowMeans(counts(dds_DGE,
    normalized = TRUE)[, dds_DGE$type == lvl]))
baseMeanPerLvl.df <- as.data.frame(baseMeanPerLvl)
baseMeanPerLvl.df$ID <- row.names(baseMeanPerLvl.df)
```

F. Export normalized counts for making count plots

```
# Export normalized counts
normalizedcounts_noIHBDt <- counts(dds_DGE, normalized = TRUE)
normalizedcounts_noIHBDt <- as.data.frame(normalizedcounts_noIHBDt)
normalizedcounts_noIHBDt$ID <- row.names(normalizedcounts_noIHBDt)
normalizedcounts_noIHBDt <- merge(collinfocoding, normalizedcounts_noIHBDt,
    by = c("ID"), all.y = TRUE)
write_delim(normalizedcounts_noIHBDt, path = "/Users/rimlandca/desktop/DeSeq2 Combined/normalizedcounts_
noIHBDt.csv" )
```

G. Results for DGE analyses between the three EHBD tissue regions

Summary function displays LFC of 0. For final results, post-hoc filtering was performed in excel for a LFC cutoff of 1.

```
# Results for GBD vs CBD tissue
resultsGtCt <- results(dds_DGE, contrast = c("type", "GBD_t", "CBD_t"), alpha = 0.05)
summary(resultsGtCt)

##
## out of 19327 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 1603, 8.3%
## LFC < 0 (down)   : 1526, 7.9%
## outliers [1]     : 19, 0.098%
## low counts [2]   : 2617, 14%
## (mean count < 3)



resultsGtCt$ID <- row.names(resultsGtCt)
GtCt.df <- as.data.frame(resultsGtCt)
GtCt.df <- merge(collinfocoding, GtCt.df, by = c("ID"), all.y = TRUE)
GtCt.df <- merge(GtCt.df, baseMeanPerLvl.df, by = c("ID"), all.y = TRUE)
write_delim(GtCt.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/GtCt_Results.csv")
```

```
# GBD vs PancD tissue
resultsGtPt <- results(dds_DGE, contrast = c("type", "GBD_t", "PancD_t"), alpha = 0.05)
summary(resultsGtPt)

##
## out of 19327 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 2416, 13%
## LFC < 0 (down)   : 2445, 13%
## outliers [1]     : 19, 0.098%
## low counts [2]   : 1873, 9.7%
## (mean count < 1)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
resultsGtPt$ID <- row.names(resultsGtPt)
GtPt.df <- as.data.frame(resultsGtPt)
GtPt.df <- merge(collinfocoding, GtPt.df, by = c("ID"), all.y = TRUE)
GtPt.df <- merge(GtPt.df, baseMeanPerLvl.df, by = c("ID"), all.y = TRUE)
write_delim(GtPt.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/GtPt_Results.csv")
```

```
# CBD vs PancD tissue
resultsCtPt <- results(dds_DGE, contrast = c("type", "CBD_t", "PancD_t"), alpha = 0.05)
summary(resultsCtPt)

##
## out of 19327 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 3056, 16%
## LFC < 0 (down)   : 3103, 16%
## outliers [1]     : 19, 0.098%
## low counts [2]   : 1499, 7.8%
## (mean count < 1)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
resultsCtPt$ID <- row.names(resultsCtPt)
CtPt.df <- as.data.frame(resultsCtPt)
CtPt.df <- merge(collinfocoding, CtPt.df, by = c("ID"), all.y = TRUE)
CtPt.df <- merge(CtPt.df, baseMeanPerLvl.df, by = c("ID"), all.y = TRUE)
write_delim(CtPt.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/CtPt_Results.csv")
```

H. Results for DGE analyses between the three EHBD organoids types

Summary function displays LFC of 0. For final results, post-hoc filtering was performed in excel for a LFC cutoff of 1.

```
# GBD vs CBD organoids
resultsGcCc <- results(dds_DGE, contrast = c("type", "GBD_c", "CBD_c"), alpha = 0.05)
summary(resultsGcCc)

##
## out of 19327 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 84, 0.43%
## LFC < 0 (down)   : 84, 0.43%
## outliers [1]     : 19, 0.098%
## low counts [2]   : 1873, 9.7%
## (mean count < 1)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
resultsGcCc.df <- as.data.frame(resultsGcCc)
resultsGcCc.df$ID <- row.names(resultsGcCc.df)

GcCc.df <- merge(collinfocoding, resultsGcCc.df, by = c("ID"), all.y = TRUE)

GcCc.df <- merge(GcCc.df, baseMeanPerLvl.df, by = c("ID"), all.y = TRUE)
write_delim(GcCc.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/GcCc_Results.csv")
```

```
# GBD cells vs PancD organoids
resultsGcPc <- results(dds_DGE, contrast = c("type", "GBD_c", "PancD_c"), alpha = 0.05)
summary(resultsGcPc)

##
## out of 19327 with nonzero total read count
## adjusted p-value < 0.05
```

```
## LFC > 0 (up)     : 216, 1.1%
## LFC < 0 (down)   : 202, 1%
## outliers [1]     : 19, 0.098%
## low counts [2]   : 2244, 12%
## (mean count < 2)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
resultsGcPc.df <- as.data.frame(resultsGcPc)
resultsGcPc.df$ID <- row.names(resultsGcPc.df)

GcPc.df <- merge(collinfocoding, resultsGcPc.df, by = c("ID"), all.y = TRUE)

GcPc.df <- merge(GcPc.df, baseMeanPerLvl.df, by = c("ID"), all.y = TRUE)
write_delim(GcPc.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/GcPc_Results.csv")
```

```
# CBD cells vs PancD organoids
resultsCcPc <- results(dds_DGE, contrast = c("type", "CBD_c", "PancD_c"), alpha = 0.05)
summary(resultsCcPc)

##
## out of 19327 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 110, 0.57%
## LFC < 0 (down)   : 129, 0.67%
## outliers [1]     : 19, 0.098%
## low counts [2]   : 1873, 9.7%
## (mean count < 1)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
resultsCcPc.df <- as.data.frame(resultsCcPc)
resultsCcPc.df$ID <- row.names(resultsCcPc.df)

CcPc.df <- merge(collinfocoding, resultsCcPc.df, by = c("ID"), all.y = TRUE)

CcPc.df <- merge(CcPc.df, baseMeanPerLvl.df, by = c("ID"), all.y = TRUE)
write_delim(CcPc.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/CcPc_Results.csv")
```

I. Results for DGE analyses between the three EHBD tissues and corresponding organoid types

Summary function displays LFC of 0. For final results, post-hoc filtering was performed in excel for a LFC cutoff of 1.

```
# CBD tissue vs CBD organoids
resultsCtCc <- results(dds_DGE, contrast = c("type", "CBD_t", "CBD_c"), alpha = 0.05)
summary(resultsCtCc)

##
## out of 19327 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 5167, 27%
## LFC < 0 (down)   : 3627, 19%
## outliers [1]     : 19, 0.098%
## low counts [2]   : 0, 0%
## (mean count < 0)
```

```
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
resultsCtCc.df <- as.data.frame(resultsCtCc)
resultsCtCc.df$ID <- row.names(resultsCtCc.df)

CtCc.df <- merge(collinfocoding, resultsCtCc.df, by = c("ID"), all.y = TRUE)

CtCc.df <- merge(CtCc.df, baseMeanPerLvl.df, by = c("ID"), all.y = TRUE)
write_delim(CtCc.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/CtCc_Results.csv")
```

```
# GBD tissue vs GBD organoids
resultsGtGc <- results(dds_DGE, contrast = c("type", "GBD_t", "GBD_c"), alpha = 0.05)
summary(resultsGtGc)

##
## out of 19327 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 3740, 19%
## LFC < 0 (down)   : 2636, 14%
## outliers [1]     : 19, 0.098%
## low counts [2]   : 1499, 7.8%
## (mean count < 1)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
resultsGtGc.df <- as.data.frame(resultsGtGc)
resultsGtGc.df$ID <- row.names(resultsGtGc.df)

GtGc.df <- merge(collinfocoding, resultsGtGc.df, by = c("ID"), all.y = TRUE)

GtGc.df <- merge(GtGc.df, baseMeanPerLvl.df, by = c("ID"), all.y = TRUE)
write_delim(GtGc.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/GtGc_Results.csv")
```

```
# PancD tissue vs PancD organoids
resultsPtPc <- results(dds_DGE, contrast = c("type", "PancD_t", "PancD_c"),
    alpha = 0.05)
summary(resultsPtPc)

##
## out of 19327 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 4229, 22%
## LFC < 0 (down)   : 3225, 17%
## outliers [1]     : 19, 0.098%
## low counts [2]   : 1499, 7.8%
## (mean count < 1)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
resultsPtPc.df <- as.data.frame(resultsPtPc)
resultsPtPc.df$ID <- row.names(resultsPtPc.df)

PtPc.df <- merge(collinfocoding, resultsPtPc.df, by = c("ID"), all.y = TRUE)

PtPc.df <- merge(PtPc.df, baseMeanPerLvl.df, by = c("ID"), all.y = TRUE)
```

```
write_delim(PtPc.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/PtPc_Results.csv")
```

J. Results of IHBD organoids compared to EHBD organoids

As all three of the extrahepatic bile duct organoids clustered closely by PCA analysis and the number of genes differentially expressed between each of the three types was minimal, we combine them into a single "EHBD organoid" group and compare the average across all three EHBD organoid types to IHBD organoids using a numeric contrast in DeSeq2.

Summary function displays LFC of 0. For final results, post-hoc filtering was performed in excel for a LFC cutoff of 1.

```
# IHBD_Huch vs all EHBD organoids
resultsImEHBDc <- results(dds_DGE, contrast = list("typeIHBD_Huch", c("typeGBD_c",
    "typeCBD_c", "typePancD_c")), listValues = c(1, -1/3), alpha = 0.05)
summary(resultsImEHBDc)
```

```
##
## out of 19327 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 1072, 5.5%
## LFC < 0 (down)   : 1158, 6%
## outliers [1]     : 19, 0.098%
## low counts [2]   : 2244, 12%
## (mean count < 2)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
resultsImEHBDc.df <- as.data.frame(resultsImEHBDc)
resultsImEHBDc.df$ID <- row.names(resultsImEHBDc.df)

resultsImEHBDc.df <- merge(collinfocoding, resultsImEHBDc.df, by = c("ID"),
    all.y = TRUE)

ImEHBDc.df <- merge(resultsImEHBDc.df, baseMeanPerLvl.df, by = c("ID"), all.y = TRUE)
write_delim(ImEHBDc.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/ImEHBDc_Results.csv")
```

```
# IHBD_ex vs all EHBD organoids

resultsIcEHBDc <- results(dds_DGE, contrast = list("typeIHBD_ex", c("typeGBD_c",
    "typeCBD_c", "typePancD_c")), listValues = c(1, -1/3), alpha = 0.05)
summary(resultsIcEHBDc)
```

```
##
## out of 19327 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 2768, 14%
## LFC < 0 (down)   : 2757, 14%
## outliers [1]     : 19, 0.098%
## low counts [2]   : 1873, 9.7%
## (mean count < 1)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
resultsIcEHBDc.df <- as.data.frame(resultsIcEHBDc)
resultsIcEHBDc.df$ID <- row.names(resultsIcEHBDc.df)
```

```
resultsIcEHBDc.df <- merge(collinfocoding, resultsIcEHBDc.df, by = c("ID"),
    all.y = TRUE)

IcEHBDc.df <- merge(resultsIcEHBDc.df, baseMeanPerLvl.df, by = c("ID"), all.y = TRUE)
write_delim(IcEHBDc.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/IcEHBDc_Results.csv")
```

```
# IHBD_ex vs IHBD_Huch organoids
resultsIcIm <- results(dds_DGE, contrast = c("type", "IHBD_ex", "IHBD_Huch"),
    alpha = 0.05)
summary(resultsIcIm)
```

```
##
## out of 19327 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 1569, 8.1%
## LFC < 0 (down)   : 1423, 7.4%
## outliers [1]     : 19, 0.098%
## low counts [2]   : 2244, 12%
## (mean count < 2)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
resultsIcIm.df <- as.data.frame(resultsIcIm)
resultsIcIm.df$ID <- row.names(resultsIcIm.df)

resultsIcIm.df <- merge(collinfocoding, resultsIcIm.df, by = c("ID"), all.y = TRUE)

IcIm.df <- merge(resultsIcIm.df, baseMeanPerLvl.df, by = c("ID"), all.y = TRUE)
write_delim(IcIm.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/IcIm_Results.csv")
```

```
# IHBD_huch vs CBD organoids
resultsImCc <- results(dds_DGE, contrast = c("type", "IHBD_Huch", "CBD_c"),
    alpha = 0.05)
summary(resultsImCc)

##
## out of 19327 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 1058, 5.5%
## LFC < 0 (down)   : 1070, 5.5%
## outliers [1]     : 19, 0.098%
## low counts [2]   : 2244, 12%
## (mean count < 2)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
resultsImCc.df <- as.data.frame(resultsImCc)
resultsImCc.df$ID <- row.names(resultsImCc.df)

resultsImCc.df <- merge(collinfocoding, resultsImCc.df, by = c("ID"), all.y = TRUE)

ImCc.df <- merge(resultsImCc.df, baseMeanPerLvl.df, by = c("ID"), all.y = TRUE)
write_delim(ImCc.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/ImCc_Results.csv")
```

```
# IHBD_huch vs GBD Organoids
```

```
resultsImGc <- results(dds_DGE, contrast = c("type", "IHBD_Huch", "GBD_c"),
    alpha = 0.05)
summary(resultsImGc)

##
## out of 19327 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 720, 3.7%
## LFC < 0 (down)   : 759, 3.9%
## outliers [1]     : 19, 0.098%
## low counts [2]   : 2617, 14%
## (mean count < 3)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
resultsImGc.df <- as.data.frame(resultsImGc)
resultsImGc.df$ID <- row.names(resultsImGc.df)

resultsImGc.df <- merge(collinfocoding, resultsImGc.df, by = c("ID"), all.y = TRUE)

ImGc.df <- merge(resultsImGc.df, baseMeanPerLvl.df, by = c("ID"), all.y = TRUE)
write_delim(ImGc.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/ImGc_Results.csv")
```

```
# IHBD_huch vs PancD Organoids
resultsImPc <- results(dds_DGE, contrast = c("type", "IHBD_Huch", "PancD_c"),
    alpha = 0.05)
summary(resultsImPc)

##
## out of 19327 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 514, 2.7%
## LFC < 0 (down)   : 584, 3%
## outliers [1]     : 19, 0.098%
## low counts [2]   : 2617, 14%
## (mean count < 3)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
resultsImPc.df <- as.data.frame(resultsImPc)
resultsImPc.df$ID <- row.names(resultsImPc.df)

resultsImPc.df <- merge(collinfocoding, resultsImPc.df, by = c("ID"), all.y = TRUE)

ImPc.df <- merge(resultsImPc.df, baseMeanPerLvl.df, by = c("ID"), all.y = TRUE)
write_delim(ImPc.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/ImPc_Results.csv")
```

III. Exploratory analysis of IHBD tissue samples

A. PCA Plots

PCA Plot of IHBD and EHBD Tissues

```
vsd_tissue_ALL <- vsd_coll_ALL[, vsd_coll_ALL$group != "cells"]
pcadata_tissue_ALL <- plotPCA(vsd_tissue_ALL, intgroup = c("type"), ntop = 5000,
    returnData = TRUE)
percentVar_tissue <- round(100 * attr(pcadata_tissue_ALL, "percentVar"))
ggplot(pcadata_tissue_ALL, aes(PC1, PC2, color = type)) + geom_point(size = 6,
    shape = 17) + xlab(paste0("PC1: ", percentVar_tissue[1], "% variance")) +
    ylab(paste0("PC2: ", percentVar_tissue[2], "% variance"))
```

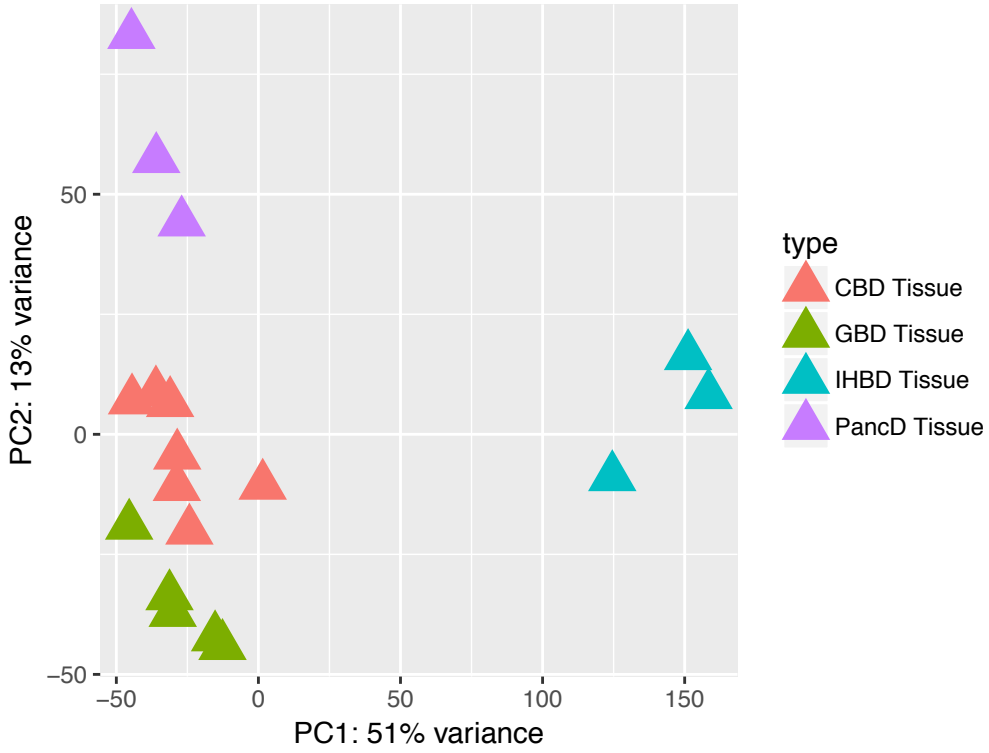


Figure 14: PCA Plot of IHBD and EHBD Tissues

PCA Plot of IHBD Tissue and IHBD Organoids

```
vsd_coll_ALL_IHBDonly <- vsd_coll_ALL[, vsd_coll_ALL$EHBD_IHBD != "EHBD" & vsd_coll_ALL$name !=
    "IHBD5_Ex"]
pcadata_IHBDonly <- plotPCA(vsd_coll_ALL_IHBDonly, intgroup = c("type", "group"),
    ntop = 5000, returnData = TRUE)
percentVar_IHBDonly <- round(100 * attr(pcadata_IHBDonly, "percentVar"))
ggplot(pcadata_IHBDonly, aes(PC1, PC2, color = type, shape = group.1)) + geom_point(size = 6) +
    xlab(paste0("PC1: ", percentVar_IHBDonly[1], "% variance")) + ylab(paste0("PC2: ",
    percentVar_IHBDonly[2], "% variance"))
```

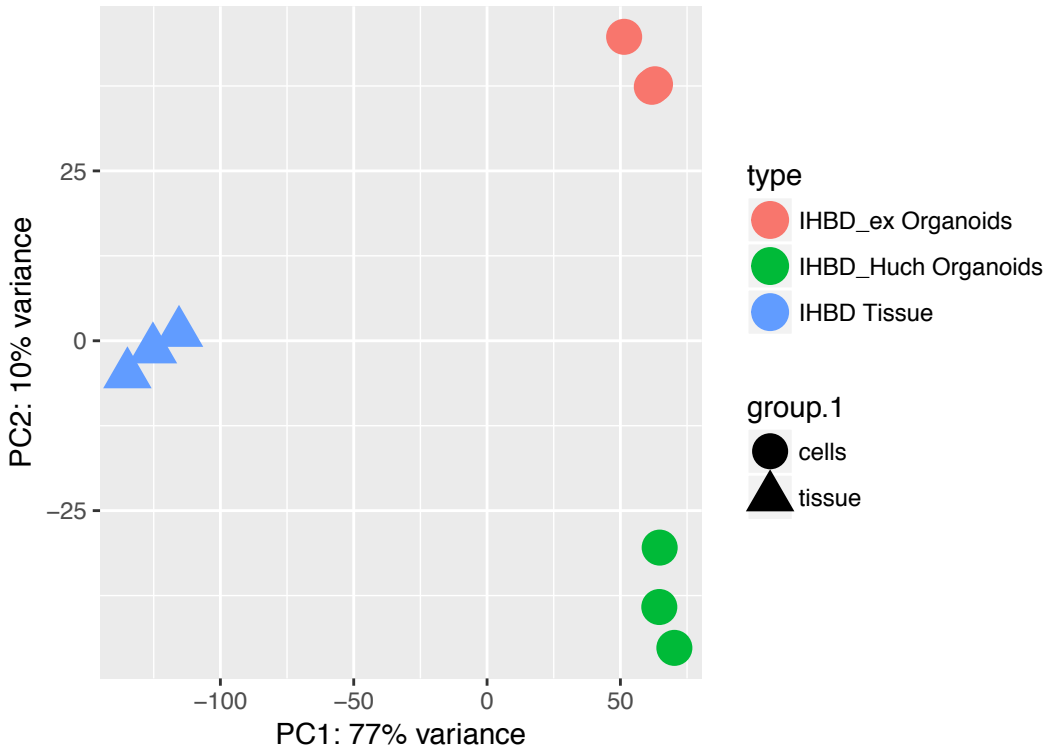


Figure 15: PCA Plot of IHBD Tissues and Organoids

B. Differential Gene Expression Analyses

Summary function displays LFC of 0. For final results, post-hoc filtering was performed in excel for a LFC cutoff of 1.

```
# DGE analyses, but now IHBD tissue in the model
ddsCollcoding_ALL <- ddsCollcoding_ALL[, ddsCollcoding_ALL$name != "IHBD5_Ex"]
ddsCollcoding_ALL$type <- droplevels(ddsCollcoding_ALL$type)
dds_DGE_ALL <- DESeq(ddsCollcoding_ALL, betaPrior = TRUE)

# Get base mean expression levels for each individual group
baseMeanPerLvl_ALL <- sapply(levels(dds_DGE_ALL$type), function(lvl) rowMeans(counts(dds_DGE_ALL,
    normalized = TRUE)[, dds_DGE_ALL$type == lvl]))
baseMeanPerLvl.df_ALL <- as.data.frame(baseMeanPerLvl_ALL)
baseMeanPerLvl.df_ALL$ID <- row.names(baseMeanPerLvl.df_ALL)

# Export normalized counts
normalizedcounts_withIHBDt <- counts(dds_DGE_ALL, normalized = TRUE)
normalizedcounts_withIHBDt <- as.data.frame(normalizedcounts_withIHBDt)
normalizedcounts_withIHBDt$ID <- row.names(normalizedcounts_withIHBDt)
normalizedcounts_withIHBDt <- merge(collinfocoding, normalizedcounts_withIHBDt,
    by = c("ID"), all.y = TRUE)
write_delim(normalizedcounts_withIHBDt, path = "/Users/rimlandca/desktop/DeSeq2 Combined/
    normalizedcounts_ withIHBDt.csv"  )
```

C. Get Results

```
# IHBD_tissue vs GBD Tissue
resultsItGt <- results(dds_DGE_ALL, contrast = c("type", "IHBD_t", "GBD_t"),
    alpha = 0.05)
summary(resultsItGt)

##
## out of 19373 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 4946, 26%
## LFC < 0 (down)   : 4458, 23%
## outliers [1]     : 21, 0.11%
## low counts [2]   : 0, 0%
## (mean count < 0)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
resultsItGt.df <- as.data.frame(resultsItGt)
resultsItGt.df$ID <- row.names(resultsItGt.df)

ItGt.df <- merge(collinfocoding, resultsItGt.df, by = c("ID"), all.y = TRUE)

ItGt.df <- merge(ItGt.df, baseMeanPerLvl.df_ALL, by = c("ID"), all.y = TRUE)
write_delim(ItGt.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/ItGt_Results.csv")
```

```
# IHBD_tissue vs CBD Tissue
resultsItCt <- results(dds_DGE_ALL, contrast = c("type", "IHBD_t", "CBD_t"),
    alpha = 0.05)
summary(resultsItCt)

##
## out of 19373 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 4989, 26%
## LFC < 0 (down)   : 4389, 23%
## outliers [1]     : 21, 0.11%
## low counts [2]   : 376, 1.9%
## (mean count < 0)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
resultsItCt.df <- as.data.frame(resultsItCt)
resultsItCt.df$ID <- row.names(resultsItCt.df)

ItCt.df <- merge(collinfocoding, resultsItCt.df, by = c("ID"), all.y = TRUE)

ItCt.df <- merge(ItCt.df, baseMeanPerLvl.df_ALL, by = c("ID"), all.y = TRUE)
write_delim(ItCt.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/ItCt_Results.csv")
```

```
# IHBD_tissue vs PancD Tissue
resultsItPt <- results(dds_DGE_ALL, contrast = c("type", "IHBD_t", "PancD_t"),
    alpha = 0.05)
summary(resultsItPt)

##
## out of 19373 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 5088, 26%
```

```
## LFC < 0 (down)   : 4440, 23%
## outliers [1]     : 21, 0.11%
## low counts [2]   : 0, 0%
## (mean count < 0)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
resultsItPt.df <- as.data.frame(resultsItPt)
resultsItPt.df$ID <- row.names(resultsItPt.df)

ItPt.df <- merge(collinfocoding, resultsItPt.df, by = c("ID"), all.y = TRUE)

ItPt.df <- merge(ItPt.df, baseMeanPerLvl.df_ALL, by = c("ID"), all.y = TRUE)
write_delim(ItPt.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/ItPt_Results.csv")
```

```
# IHBD_tissue vs EHBD tissues
resultsItEt <- results(dds_DGE_ALL, contrast = list("typeIHBD_t", c("typeGBD_t",
    "typeCBD_t", "typePancD_t")), listValues = c(1, -1/3), alpha = 0.05)

summary(resultsItEt)
```

```
##
## out of 19373 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 5363, 28%
## LFC < 0 (down)   : 4759, 25%
## outliers [1]     : 21, 0.11%
## low counts [2]   : 376, 1.9%
## (mean count < 0)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
resultsItEt.df <- as.data.frame(resultsItEt)
resultsItEt.df$ID <- row.names(resultsItEt.df)

ItEt.df <- merge(collinfocoding, resultsItEt.df, by = c("ID"), all.y = TRUE)

ItEt.df <- merge(ItEt.df, baseMeanPerLvl.df_ALL, by = c("ID"), all.y = TRUE)
write_delim(ItEt.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/ItEt_Results.csv")
```

```
# IHBD_tissue vs IHBD_Huch Organoids
resultsItIm <- results(dds_DGE_ALL, contrast = c("type", "IHBD_t", "IHBD_Huch"),
    alpha = 0.05)
summary(resultsItIm)
```

```
##
## out of 19373 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 5563, 29%
## LFC < 0 (down)   : 3836, 20%
## outliers [1]     : 21, 0.11%
## low counts [2]   : 376, 1.9%
## (mean count < 0)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
resultsItIm.df <- as.data.frame(resultsItIm)
resultsItIm.df$ID <- row.names(resultsItIm.df)

ItIm.df <- merge(collinfocoding, resultsItIm.df, by = c("ID"), all.y = TRUE)

ItIm.df <- merge(ItIm.df, baseMeanPerLvl.df_ALL, by = c("ID"), all.y = TRUE)
write_delim(ItIm.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/ItIm_Results.csv")
```

```
# IHBD_tissue vs IHBD_ex Organoids
resultsItIc <- results(dds_DGE_ALL, contrast = c("type", "IHBD_t", "IHBD_ex"),
    alpha = 0.05)
summary(resultsItIc)

##
## out of 19373 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 5933, 31%
## LFC < 0 (down)   : 3975, 21%
## outliers [1]     : 21, 0.11%
## low counts [2]   : 376, 1.9%
## (mean count < 0)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results
```

```
resultsItIc.df <- as.data.frame(resultsItIc)
resultsItIc.df$ID <- row.names(resultsItIc.df)

ItIc.df <- merge(collinfocoding, resultsItIc.df, by = c("ID"), all.y = TRUE)

ItIc.df <- merge(ItIc.df, baseMeanPerLvl.df_ALL, by = c("ID"), all.y = TRUE)
write_delim(ItIc.df, path = "/Users/rimlandca/desktop/DeSeq2 Combined/ItIc_Results.csv")
```